

A;Accession: A27441

A:Molecule type: mRNA
A:Residues: 1-713 <LA2>
R:Lapeyre, B.; Amalric, F.; Ghaffari, S.H.; Venkatarama Rao, S.V.; Dumar, T.S.; Olson, J. Biol. Chem. 261, 9167-9173, 1986
A:Title: Protein and cDNA sequence of a glycine-rich, dimethylarginine-containing region
A:Reference number: A24808; MUID:86250857; PMID:3755137
A:Accession: A24808
A:Molecule type: mRNA
A:Residues: 522-541, 'G', 543-544, 'R', 546-713 <LA2>
A:Cross-references: GB:M13757; NID:g191004; PIDN:AAA36961.1; PID:g387047
R:Lapeyre, B.; Caillergues-Ferrer, M.; Bouche, G.; Amalric, F.
Nucleic Acids Res. 13, 5805-5816, 1985
A:Title: Cloning of cDNA encoding a 100 kDa nucleolar protein (nucleoline) of Chinese hamster ovary cells
A:Reference number: A24070; MUID:85297777; PMID:2994013
A:Accession: A24070
A:Molecule type: mRNA
A:Residues: 179-238 <LA3>
C:Superfamily: nucleolin; ribonucleoprotein repeat homology
C:Keywords: DNA binding; nucleus; phosphoprotein
C:Accession: A24070
A:Title: Sequence and analysis of the plant Arabidopsis thaliana
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: D85087
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-715 <STO>
C:Cross-references: GB:NC_001268; NID:g7267513; PIDN:CAB77996.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g08710
A:Map position: 4

Query Match 38.7%; Score 79; DB 2; Length 715;
Best Local Similarity 42.9%; Pred. No. 5.8;
Matches 15; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 7 VSSKDKENSKNDVLDKEKEAEETEEELKEK 41
DB 74 LTEQDPENVEESBEKEKEKEEKEEGEE 108

RESULT 4
D81798
lactoferrin-binding protein NMA1740 [imported] - Neisseria meningitidis (strain Z2491 se
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
A:Accession: D81798
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: D81798

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-741 <PAR>
A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB84968.1; PID:g7380382
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: lbpB; NMA1740

Query Match 38.7%; Score 79; DB 2; Length 741;
Best Local Similarity 51.5%; Pred. No. 6.1;
Matches 17; Conservative 8; Mismatches 6; Indels 2; Gaps 1;

QY 8 SSKDKENSKNDVLDKEKEAEETEEELKEK 40
DB 510 NSEDEEETAEEDDD--EAEEEVEEPESPEE 540

RESULT 5
A35466
progesterone receptor form B - Chicken
N:Contains: progesterone receptor form A
C:Species: Gallus gallus (chicken)
C:Date: 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
A:Accession: A35466; S06284; A40903; A24661; A24312; A40911; A61552
R:Jeltsch, J.M.; Turcotte, B.; Garnier, J.M.; Lerouge, T.; Krozowski, Z.; Gronemeyer, H.;
J. Biol. Chem. 265, 3967-3974, 1990
A:Title: Characterization of multiple mRNAs originating from the chicken progesterone re
A:Reference number: A35466; MUID:90154085; PMID:2303488
A:Accession: A35466
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-786 <JEL>
A:Cross-references: GB:M32732; GB:J05240; NID:g212554; PIDN:AAA49011.1; PID:g212558
R:Gronemeyer, H.; Turcotte, B.; Quirin-Stricker, C.; Bocquel, M.T.; Meyer, M.E.; Krozowski
EMBO J. 6, 3985-3994, 1987
A:Title: The chicken progesterone receptor: sequence, expression and functional analysis.
A:Reference number: S06284; MUID:88166640; PMID:3443098
A:Accession: S06284
A:Molecule type: DNA
A:Residues: 1-786 <GRO>
A:Cross-references: EMBL:Y00092; NID:g63744; PIDN:CAA68282.1; PID:g63745
R:Conneely, O.M.; Dobson, A.D.W.; Tsai, M.J.; Beattie, W.G.; Toft, D.O.; Huckaby, C.S.;
Mol. Endocrinol. 1, 517-525, 1987
A:Title: Sequence and expression of a functional chicken progesterone receptor.
A:Reference number: A40903; MUID:91042592; PMID:3153474
A:Accession: A40903
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-64, 'E', 65-786 <CON>
A:Cross-references: GB:M37518
R:Conneely, O.M.; Sullivan, W.P.; Toft, D.O.; Birnbaumer, M.; Cook, R.G.; Maxwell, B.L.;
Science 233, 767-770, 1986
A:Title: Molecular cloning of the chicken progesterone receptor.
A:Reference number: A24661; MUID:86289413; PMID:2426779
A:Accession: A24661
A:Molecule type: mRNA
A:Residues: 128-133, 'E', 135-147, 'E', 149-164 <CO2>
A:Note: the authors translated the codon CAG for residue 7 as Glu and CAG for residue 21
R:Jeltsch, J.M.; Krozowski, Z.; Quirin-Stricker, C.; Gronemeyer, H.; Simpson, R.J.; Garni
Proc. Natl. Acad. Sci. U.S.A. 83, 5424-5428, 1986
A:Title: Cloning of the chicken progesterone receptor.
A:Reference number: A24312; MUID:86287271; PMID:2426697
A:Accession: A24312
A:Molecule type: mRNA
A:Residues: 417-490 <JE2>
A:Cross-references: GB:M14280; NID:g212607; PIDN:AAA49039.1; PID:g212608
A:Note: amino acid and corresponding nucleotide sequences are also shown for three small
R:Birnbaumer, M.; Hinrichs-Rosello, M.V.; Cook, R.G.; Schrader, W.T.; O'Malley, B.W.
Mol. Endocrinol. 1, 249-259, 1987
A:Title: Chemical and antigenic properties of pure 108,000 molecular weight chick prostag
A:Reference number: A40911; MUID:88288199; PMID:3453892
A:Accession: A40911
A:Status: preliminary

A;Molecule type: protein
 A;Residues: 128-133, 'E', 135-147, 'E', 149-164, 546-558 <BIR>
 R;Simpson, R.J.; Grego, B.; Govindan, M.V.; Gronemeyer, H.
 Mol. Cell. Endocrinol. 52, 177-184, 1987
 A;Title: Peptide sequencing of the chick oviduct progesterone receptor form B.
 A;Reference number: A61552; MUID:88005426; PMID:363503
 A;Accession: A61552
 A;Molecule type: protein
 A;Residues: 136-153, 168-174, 195-228, 526-537, 'X', 539, 546-563 <SIM>
 C;Genetics:
 A;Introns: 400/2; 451/1; 490/1; 591/1; 639/2; 683/1; 735/3
 C;Superfamily: progesterone receptor; erba transforming protein homology
 C;Keywords: DNA binding; nucleus; steroid hormone receptor; zinc finger
 F;1-786/Product: progesterone receptor form B #status predicted <MAI>
 F;128-786/Product: progesterone receptor form A #status predicted <MA2>
 F;419-682/Domain: erba transforming protein homology <ERBA>
 F;421-441/Region: zinc finger
 F;457-481/Region: zinc finger

Query Match 38.7%; Score 79; DB 2; Length 786;
 Best Local Similarity 45.9%; Pred. No. 6.4;
 Matches 17; Conservative 9; Mismatches 9; Indels 2; Gaps 1;
 Qy 5 LYVSKDKENISKENDVLDKEKEAEETEELEEK 41
 Db 42 LYPRSDDEEEENEB--EEEEEPOQREEEEEEE 76
 C;Genetics:
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-2924 <BAR>
 A;Cross-references: EMBL:U27338; NID:9914918; PID:9914919; PIDN:AAB60251.1
 C;Genetics:
 A;Gene: EMP1
 A;Introns: 2476/3

RESULT 6
 T18378
 variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
 N;Alternate names: erythrocyte membrane binding protein 1 (EMP1)
 C;Species: Plasmodium falciparum
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C;Accession: T18378
 R;Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi, T.
 Cell 82, 77-87, 1995
 A;Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and
 A;Reference number: Z18925; MUID:95330812; PMID:7541722
 A;Accession: T18378
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-2924 <BAR>
 A;Cross-references: EMBL:U27338; NID:9914918; PID:9914919; PIDN:AAB60251.1
 C;Genetics:
 A;Gene: EMP1
 A;Introns: 2476/3

Query Match 38.7%; Score 79; DB 2; Length 2924;
 Best Local Similarity 45.5%; Pred. No. 23;
 Matches 15; Conservative 9; Mismatches 9; Indels 0; Gaps 0;
 Qy 8 SSKDKENISKENDVLDKEKEAEETEELEEK 40
 Db 777 SDDDDFFEEEEEDEGEAEAEVQEKTE 809
 C;Genetics:
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-231 <SCH>
 A;Cross-references: EMBL:ALJ56815; GSPDB:GN001116; NCSP:B24H17.160
 A;Experimental source: BAC clone B24H17; strain OR74A
 C;Genetics:

RESULT 7
 T49830
 hypothetical protein B24H17.160 [imported] - Neurospora crassa
 C;Species: Neurospora crassa
 C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C;Accession: T49830
 R;Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, May 2000
 A;Reference number: Z25022
 A;Accession: T49830
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-231 <SCH>
 A;Cross-references: EMBL:ALJ56815; GSPDB:GN001116; NCSP:B24H17.160
 A;Experimental source: BAC clone B24H17; strain OR74A
 C;Genetics:

A;Gene: NCSP:B24H17.160
 A;Map position: 6
 A;Introns: 38/1; 102/3; 128/3

Query Match 38.2%; Score 78; DB 2; Length 231;
 Best Local Similarity 48.5%; Pred. No. 2.4;
 Matches 16; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy 9 SKDKENISKENDVLDKEKEAEETEELEEK 41
 Db 44 SKDEEEEEEDEEEEEEDEEEEEEDEEEEEE 76
 C;Genetics:
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-387 <BUL>
 C;Genetics:
 A;Gene: FlyBase:up
 A;Cross-references: FlyBase:FBgn0004169
 C;Keywords: muscle

RESULT 8

S02708
 troponin T - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster
 C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 16-Feb-1997
 C;Accession: S02708
 R;Bullard, B.; Leonard, K.; Larkins, A.; Butcher, G.; Karlik, C.; Fyrberg, E.
 J. Mol. Biol. 204, 621-637, 1988
 A;Title: Troponin of asynchronous flight muscle.
 A;Reference number: S02708; MUID:89141761; PMID:2852258
 A;Accession: S02708
 A;Status: nucleic acid sequence not shown; not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-387 <BUL>
 C;Genetics:
 A;Gene: FlyBase:up
 A;Cross-references: FlyBase:FBgn0004169
 C;Keywords: muscle

Query Match 37.7%; Score 77; DB 2; Length 387;
 Best Local Similarity 45.7%; Pred. No. 4.9;
 Matches 16; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

Qy 7 VSSKDKENISKENDVLDKEKEAEETEELEEK 41
 Db 347 VEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 381
 C;Genetics:
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-2924 <BAR>
 A;Cross-references: EMBL:U27338; NID:9914918; PID:9914919; PIDN:AAB60251.1
 C;Genetics:
 A;Gene: EMP1
 A;Introns: 2476/3

RESULT 9

S13251
 troponin T - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster
 C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Sep-1998
 C;Accession: S13251; S29482
 R;Fyrberg, E.; Fyrberg, C.C.; Beall, C.; Saville, D.L.
 J. Mol. Biol. 216, 657-675, 1990
 A;Title: Drosophila melanogaster troponin-T mutations engender three distinct syndromes
 A;Reference number: S13251; MUID:91080155; PMID:2124273
 A;Accession: S13251
 A;Molecule type: mRNA
 A;Residues: 1-396 <FYR1>
 A;Cross-references: EMBL:X54504
 R;Fyrberg, E.A.
 submitted to the EMBL Data Library, August 1990
 A;Reference number: S29482
 A;Accession: S29482
 A;Molecule type: mRNA
 A;Residues: 1-191, 'A', 193-396 <FYR2>
 A;Cross-references: EMBL:X54504; NID:98739; PID:98740
 C;Genetics:
 A;Gene: FlyBase:up
 A;Cross-references: FlyBase:FBgn0004169
 A;Map position: X

Query Match 37.7%; Score 77; DB 2; Length 396;
 Best Local Similarity 45.7%; Pred. No. 5;
 Matches 16; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

Qy 7 VSSKDKENISKENDVLDKEKEAEETEELEEK 41
 Db 347 VEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 381
 C;Genetics:
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-231 <SCH>
 A;Cross-references: EMBL:ALJ56815; GSPDB:GN001116; NCSP:B24H17.160
 A;Experimental source: BAC clone B24H17; strain OR74A
 C;Genetics:

Db 356 VEEDEDEDEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE 390

RESULT 10
A54514
Glutamic acid-rich protein precursor - malaria parasite (Plasmodium falciparum)
N;Alternate names: GARP
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 21-Jul-2000
C;Accession: A54514
R;Triglia, T.; Stahl, H.D.; Crewther, P.E.; Silva, A.; Anders, R.F.; Kemp, D.J.
Mol. Biochem. Parasitol. 31, 199-202, 1988
A;Title: Structure of a Plasmodium falciparum gene that encodes a glutamic acid-rich pr
A;Reference number: A54514; MUID:89040048; PMID:2903445
A;Accession: A54514
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-678 <TRI>
A;Cross-references: GB:J03998; NID:g160298; PIDN:AAA29605.1; PID:g160299
C;Keywords: tandem repeat

A;Introns: 25/3
C;Superfamily: histone H1
C;Keywords: tandem repeat

Query Match 37.7%; Score 77; DB 2; Length 678;
Best Local Similarity 43.8%; Pred.No. 8.4;
Matches 14; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 10 KXENISKENDVDLDEKEEAETETELEEK 41
:|:: : : :||| |||||
Db 567 EDEEVEDEEEEEEEEEEEEEEEEEEEEE 598

RESULT 11
S29795
Hypothetical protein 2280 - evening primrose (Oenothera picensis subsp. picensis) chloro
C;Species: chloroplast Oenothera picensis subsp. picensis (evening primrose)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: S29795; S15983
R;Nimzyk, R.; Schoendorf, T.; Hachtel, W.
Curr. Genet. 23, 265-270, 1993
A;Title: In-frame length mutations associated with short tandem repeats are located in u
A;Reference number: S29795; MUID:93169690; PMID:8435856
A;Accession: S29795
A:Molecule type: DNA
A;Residues: 1-721 <NIM>
A;Cross-references: EMBL:X64616; NID:g14334; PID:g14335
C;Keywords: chloroplast

Query Match 37.7%; Score 77; DB 2; Length 721;
Best Local Similarity 51.6%; Pred.No. 8.9;
Matches 16; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 11 DKENISKENDVDLDEKEEAETETELEEK 41
:|:: : : :||| |||||
Db 555 EEEEPKEEDLHEEEEEEEEEEEEEEEEE 585

RESULT 12
S65169
Hypothetical protein YPL158c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein P2570
C;Species: Saccharomyces cerevisiae
C;Date: 10-Dec-1994 #sequence_revision 31-May-1996 #text_change 19-Apr-2002
C;Accession: S65169; S69441
R;Purnelle, B.; Coster, F.; Goffeau, A.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S65154
A;Accession: S65169
A:Molecule type: DNA
A;Residues: 1-758 <PUR>

QY 7 VSSXDKENISKEN-----DDVLDEKEEAEETEELEEK 41
 Db 125 VDREDEBEIEEEDDDDDGDDVVEEEDDDDEEEEEE 169

RESULT 15

T46608
 zinc finger protein Png-1 - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
 C/Accession: T46608
 R/Weiner, J.; Chun, J.
 J. Comp. Neurol. 381, 130-142, 1997
 A/Title: Png-1, a nervous system-specific zinc finger gene, identifies regions containin
 A/Reference number: 223102; MUID:97276971; PMID:9130664
 A/Accession: T46608
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-1188 <WEI>
 A/Cross-references: EMBL:U86338; NID:gl835754; PID:gl835755; PIDN:AACS3157.1
 A/Experimental source: strain BALB/c
 A/Genetics:
 A/Name: Png-1

Query Match 37.7%; Score 77; DB 2; Length 1188;
 Best Local Similarity 40.9%; Pred. No. 14;
 Matches 18; Conservative 9; Mismatches 7; Indels 10; Gaps 1;

QY 8 SSKDKENISKEN-----DDVLDEKEEAEETEELEEK 41
 Db 128 SREDEBEIEEEDDDDDGDDVVEEEDDDDEEEEEE 171

Search completed: September 22, 2004, 18:10:02
 Job time : 16.5868 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:32 ; Search time 8.47107 Seconds
(without alignments)
252.020 Million cell updates/sec

Title: US-10-774-602-14
Perfect score: 204
Sequence: 1 MLHLVSVKXENISKEND.....VLDEKEERAEETEEBLEEK 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	84	41.2	1121	1 MYT1 HUMAN	Q01538 homo sapien
2	82	40.2	3135	1 S230 PLAFO	Q08372 plasmodium
3	81	39.7	120	1 NUCL ARATH	Q01et3 arabidopsis
4	79	38.7	713	1 NUCLE MESAU	P08199 mesocricetu
5	79	38.7	786	1 PRGR CHICK	P07812 gallus gall
6	77.5	38.0	627	1 BALA XENLA	Q08vr5 xenopus lae
7	77	37.7	396	1 TRT DRONE	P19351 drosophila
8	77	37.7	678	1 GARE PLAFF	P13816 plasmodium
9	77	37.7	721	1 YCF2 OENPI	P31568 oenothera p
10	76.5	37.5	758	1 YP58 YEAST	Q99299 saccharomyc
11	76.5	37.5	1124	1 TCF8 HUMAN	P37275 homo sapien
12	76	37.3	2492	1 ATRX HUMAN	P46100 homo sapien
13	76	37.3	2492	1 ATRX PANTR	Q7yqm4 pan troglod
14	76	37.3	2492	1 ATRX PONPY	Q7yqm3 pongo pygma
15	75.5	37.0	213	1 NEUM CARAU	P17691 carassius a
16	75.5	37.0	1257	1 RBBI HUMAN	P29374 homo sapien
17	75	36.8	7039	1 DAXX MOUSE	Q35613 mus muscucu
18	75	36.8	1043	1 TCF8 MESAU	Q60542 mesocricetu
19	75	36.8	1099	1 NKX1 HUMAN	Q60721 homo sapien
20	75	36.8	1263	1 BAT6 MOUSE	Q92148 mus muscucu
21	74	36.3	180	1 HMGI CRIGR	P07156 cricetus
22	74	36.3	214	1 HMGI MOUSE	P07155 mus muscucu
23	74	36.3	301	1 TRT2 CHICK	P02642 gallus gall
24	74	36.3	417	1 CRT2 BOVIN	P52193 bos taurus
25	74	36.3	421	1 CRT2 BOVIN	P42918 bos taurus
26	74	36.3	630	1 YCF2 OENVI	P31569 oenothera v
27	74	36.3	1210	1 BAT8 HUMAN	Q96kq7 homo sapien
28	73.5	36.0	202	1 IPPD BOVIN	P07516 bos taurus
29	73.5	36.0	321	1 ABRA PLAFF	P23746 plasmodium
30	73	35.8	110	1 THYA HUMAN	P06454 homo sapien
31	73	35.8	110	1 THYA MOUSE	P26350 mus muscucu
32	73	35.8	111	1 THYA RAT	P06302 rattus norv
33	73	35.8	137	1 IPPD FIG	Q29277 sus scrofa

RESULT 1																	
MYT1_HUMAN	73	35.8	204	1	IPPD_HUMAN	73	35.8	399	1								
ID	Q01538	Q94922	Q9UPV2		STANDARD												
AC	01-JUL-1993	(Rel. 26, Created)			PRT				1121 AA.								
DT	16-OCT-2001	(Rel. 40, Last sequence update)															
DT	15-MAR-2004	(Rel. 43, Last annotation update)															
DE	Myelin transcription factor 1 (MYT1) (MYT1) (Proteolipid protein binding protein) (PLPB1).																
DE	binding protein) (PLPB1).																
GN	MYT1 OR PLPB1 OR MYT1 OR MYT1 OR KIAA0835 OR KIAA1050.																
OS	Homo sapiens (Human).																
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.																
OX	NCBI_TaxID=9606;																
RN	[1]																
RP	SEQUENCE FROM N.A.																
RC	TISSUE=Brain;																
RX	MEDLINE=99156230; PubMed=10048485;																
RA	Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirotsawa M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;																
RA	"Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."																
RL	DNA Res. 5:355-364(1998).																
RN	[2]																
RP	SEQUENCE FROM N.A.																
RX	MEDLINE=21638749; PubMed=11780052;																
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clegg S., Cobley V.E., Collier R.E., Clark L.N., Clark S.Y., Clee C.M., Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Hunt A.R., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hurley J.L., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Leharvalho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.B., McAninchie L.J., Mcay K., McMurray A.A., Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E., Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.;																
RT	The DNA sequence and comparative analysis of human chromosome 20.																
RL	Nature 414:865-871(2001).																
RN	[3]																
RP	SEQUENCE OF 293-1121 FROM N.A.																

Q9ud71 homo sapien
P49709 carassius a
Q9y02 rattus norv
P32892 saccharomyc
Q28139 bos taurus
Q60840 homo sapien
Q62947 rattus norv
P10103 bos taurus
P12682 sus scrofa
P34618 caenorhabdi
Q94a89 arabidopsis
P13383 rattus norv

[1]


```
FT DOMAIN 1 420 MODULATING, PRO-RICH.
FT DNA_BIND 421 486 NUCLEAR RECEPTOR-TYPE.
FT C4-TYPE.
FT ZN_FING 421 441 C4-TYPE.
FT PROSITE; PS00633; BROMODOMAIN 1; 1.
FT DOMAIN 457 481 STEROID-BINDING.
FT DOMAIN 487 786 ASP/GLU-RICH (ACIDIC).
FT VARSPLIC 48 80 Missing (in isoform B and isoform B').
FT VARSPLIC 1 127 /FTID=VSP_003707.
FT VARSPLIC 452 458 QHNYLCA -> TISYHCS (in isoform A' and isoform B').
FT VARSPLIC 459 786 /FTID=VSP_003708.
FT VARSPLIC 459 786 Missing (in isoform A' and isoform B').
FT CONFLICT 58 58 E -> DD (IN REF. 2).
FT CONFLICT 480 480 K -> N (IN REF. 2).
FT CONFLICT 489 489 G -> A (IN REF. 2).
FT CONFLICT 577 577 R -> T (IN REF. 2).
FT CONFLICT 642 642 M -> I (IN REF. 2).
FT CONFLICT 786 AA; 85743 MW; 659559950BC45ED9 CRC64;
SQ SEQUENCE 786 AA; 85743 MW; 659559950BC45ED9 CRC64;

Query Match 38.7%; Score 79; DB 1; Length 786;
Best Local Similarity 45.9%; Pred. No. 4.9;
Matches 17; Conservative 9; Mismatches 9; Indels 2; Gaps 1;

QY 5 LYVSSKDKENISKENDVLDKEKEEAETEEELK 41
Db 42 LYPRSDDEEEENEE--EEEEEPPQREEEEEE 76

RESULT 6
ID BAIA_XENLA STANDARD; PRT; 627 AA.
AC Q8UVR5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ATP-utilizing chromatin assembly and remodel factor 1 (xACF1) (Fragment).
DE (Fragment).
GN ACF1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21668941; PubMed=11809820;
RA MacCallum D.E., Losada A., Kobayashi R., Hirano T.;
RA "ISWI remodeling complexes in Xenopus egg extracts: identification as major chromosomal components that are regulated by INCENP-aurora B."; Mol. Biol. Cell 13:25-39(2002).
CC -1- FUNCTION: May play a role in transcriptional regulation.
CC -1- SUBUNIT: Together with p18 and p20 proteins, it forms the xenopus version of CHRAC.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: Phosphorylated in mitosis.
CC -1- SIMILARITY: Belongs to the WAL family.
CC -1- SIMILARITY: Contains 1 bromodomain.
CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF412332; AAL60160.1; -
CC InterPro; IPR001487; Bromodomain.
CC InterPro; IPR001965; Znf PHD.
CC Pfam; PF00439; bromodomain; 1.
CC Pfam; PF00628; PHD; 1.
```

```
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 1.
DR SMART; SM00249; PHD; 1.
DR PROSITE; PS00633; BROMODOMAIN 1; 1.
DR PROSITE; PS00014; BROMODOMAIN 2; 1.
DR PROSITE; PS01359; ZF PHD 1; 1.
DR PROSITE; PS00016; ZF PHD 2; 1.
DR Transcription regulation; Bromodomain; Zinc-finger; Coiled coil;
KW Nuclear protein; Phosphorylation.
FT NON_TER 1 1
FT ZN_FING 222 272 PHD-TYPE.
FT DOMAIN 518 588 BROMODOMAIN.
FT DOMAIN 281 327 COILED COIL (POTENTIAL).
SQ SEQUENCE 627 AA; 70208 MW; 42C17587E531D380 CRC64;

Query Match 38.0%; Score 77.5; DB 1; Length 627;
Best Local Similarity 45.9%; Pred. No. 5.3;
Matches 17; Conservative 5; Mismatches 10; Indels 5; Gaps 1;

QY 4 HLYVSSKDKENISKENDVLDKEKEEAETEEELK 40
Db 282 HRYSDMSD-----EEEEELDKKEEEEEEQEELSE 313

RESULT 7
ID TRT_DROME STANDARD; PRT; 396 AA.
AC PI9351;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Troponin T, skeletal muscle (Upheled protein) (Intended thorax DE protein).
GN UP OR INT.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Asynchronous muscle;
RX MEDLINE=89141761; PubMed=28522258;
RA Bullard B., Leonard K., Larkins A., Butcher G., Karlik C.,
RA Fryberg E.A.;
RT "Troponin of asynchronous flight muscle.";
RL J. Mol. Biol. 204:621-637(1988).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RX STRAIN=Canton-S;
RX MEDLINE=91080155; PubMed=2124273;
RA Fryberg E.A., Fryberg C.C., Beall C., Saville D.L.;
RT "Drosophila melanogaster troponin-T mutations engender three distinct syndromes of myofibrillar abnormalities.";
RL J. Mol. Biol. 216:657-675(1990).
CC -1- FUNCTION: Troponin T is the tropomyosin-binding subunit of troponin, the thin filament regulatory complex which confers calcium-sensitivity to striated muscle actomyosin ATPase activity.
CC Mutations in troponin T engender three distinct syndromes of myofibrillar abnormalities.
CC -1- SIMILARITY: Belongs to the troponin T family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X54504; CAA38366.1; -
CC FIR; S13251; S13251.
CC FlyBase; FBgn0004169; up.
```



```

DR InterPro; IPR001978; Troponin.
DR Pfam; PF00992; Troponin; 1.
KW Muscle protein.
FT DOMAIN 341 396 ASP/GLU-RICH (HIGHLY ACIDIC).
SQ SEQUENCE 396 AA; 47333 MW; 3CAAA8AA8943C23 CRC64;

Query Match 37.7%; Score 77; DB 1; Length 396;
Best Local Similarity 45.7%; Pred. No. 3.8;
Matches 16; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 7 VSKDKENISKENDVLDKEEAEETEELEEK 41
   ||||| : : : : : ||||| : : : : :
Db 356 VEEDEEEDDEEEDEEEDEEEDEEEDEEE 390

RESULT 8
GARP PLAFF
ID GARP PLAFF STANDARD; PRT; 678 AA.
KW P13816;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Glutamic acid-rich protein precursor.
GN GARP.
OS Plasmodium falciparum (isolate FC27 / Papua New Guinea).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5837;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89040048; PubMed=2903445;
RA Triglia T., Stahl H.-D., Crewther P.E., Silva A., Anders R.F.,
RA Kemp D.J.;
RT "Structure of a Plasmodium falciparum gene that encodes a glutamic
  acid-rich protein (GARP).";
RL Mol. Biochem. Parasitol. 31:199-202(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03998; AAA29605.1; -.
DR PIR; A54514; A54514.
RE Repeat; Malaria; Antigen; Signal.
KW SIGNAL 1 25
FT CHAIN 26 678 GLUTAMIC ACID-RICH PROTEIN.
FT DOMAIN 120 164 15 X 3 AA TANDEM REPEATS OF K-K-X.
FT DOMAIN 372 416 9 X APPROXIMATE TANDEM REPEATS.
FT DOMAIN 417 441 5 X APPROXIMATE TANDEM REPEATS.
FT DOMAIN 576 604 POLY-GLU.
FT DOMAIN 605 653 7 X APPROXIMATE TANDEM REPEATS.
FT DOMAIN 654 663 POLY-GLU.
SQ SEQUENCE 678 AA; 80551 MW; 2A8F85606496EA9E CRC64;

Query Match 37.7%; Score 77; DB 1; Length 678;
Best Local Similarity 43.8%; Pred. No. 6.4;
Matches 14; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 10 KDKENISKENDVLDKEEAEETEELEEK 41
   ||||| : : : : : ||||| : : : : :
Db 567 EDEEVEEDEEEDEEEDEEEDEEEDEEE 598

RESULT 9
YCF2 OENPI
ID YCF2 OENPI STANDARD; PRT; 721 AA.
AC P31568;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Hypothetical protein ycf2 (ORF 2280) (Fragment).
GN YCF2.
OS Oenothera picensis (Oenothera odorata).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
  Myrtales; Onagraceae; Oenothera.
OX NCBI_TaxID=3946;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93169690; PubMed=8435856;
RA Nimzyk R., Shoendorf T., Hachtel W.;
RT "In-frame length mutations associated with short tandem repeats are
  located in unassigned open reading frames of Oenothera chloroplast
  DNA.";
RL Curr. Genet. 23:265-270(1993).
CC -----
CC -!- FUNCTION: Not yet known.
CC -!- SIMILARITY: Belongs to the ycf2 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X64616; CAA45898.1; -.
DR PIR; S29795; S29795.
KW Chloroplast; Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 721 AA; 82900 MW; 279A3FC1AA06F9C9 CRC64;

Query Match 37.7%; Score 77; DB 1; Length 721;
Best Local Similarity 51.6%; Pred. No. 6.7;
Matches 16; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 11 DKENISKENDVLDKEEAEETEELEEK 41
   ::||| : : : : : ||||| : : : : :
Db 555 EEEEPKKEEDELHEEEDEEEDEEEDEEE 585

RESULT 10
YPS8 YEAST
ID YPS8 YEAST STANDARD; PRT; 758 AA.
AC Q99299;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein YPL158C.
GN YPL158C OR P2370.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
  Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / AB972;
RX MEDLINE=97103777; PubMed=8948103;
RA Purnelle B., Coster F., Goffeau A.;
RT "The sequence of 55 kb on the left arm of yeast chromosome XVI
  identifies a small nuclear RNA, a new putative protein kinase and two
  new putative regulators.";
RL Yeast 12:1483-1492(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / AB972;
RX MEDLINE=97313271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,
  Araujo R., Aparicio A., Barrall B.G., Badcock K., Benes V.,
  Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
  Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
  Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,

```

RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W.,
RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
RA Komp C., Kuri O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
RA Urrutkarazu L.A., Ushinsky S., Vlerendeels F., Viessers S., Voss H.,
RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA Zhong W.W., Zollner A., Vo D.H., Hani J.,
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI";
RL Nature 387:103-105(1997).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X96770; CAA65563.1; -;
CC EMBL; Z73514; CAA97863.1; -;
CC EMBL; S65169; S65169.
CC GERMONLINE; 141140; -;
CC SGD; S0006079; YPL158C.
CC KW Hypothetical protein; Coiled coil.
CC FT DOMAIN 253 263 POLY-SER.
CC FT DOMAIN 648 711 COILED COIL (POTENTIAL).
CC FT DOMAIN 652 658 POLY-GLU.
CC FT DOMAIN 664 669 POLY-GLU.
CC FT DOMAIN 686 690 POLY-GLU.
CC SQ SEQUENCE 758 AA; 84845 MW; 384FA92B91C87F2B CRC64;
Query Match 37.7%; Score 77; DB 1; Length 758;
Best Local Similarity 46.9%; Pred. No. 7.1;
Matches 15; Conservative 8; Mismatches 9; Indels 0; Gaps 0;
QY 10 KDKENISKENDVDLDEKEEAETEELEEK 41
DB 661 EDDEEEEDDEDEDEDEDEDEDEDEDEEKR 692
RESULT 11
TCF8_HUMAN
ID TCF8_HUMAN STANDARD; PRT; 1124 AA.
AC P37275; Q12924; Q13800;
DA 01-OCT-1994 (Rel. 30, Created)
DE 15-JUL-1999 (Rel. 38, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcription factor 8 (NTL-2-A zinc finger protein) (Negative
DE regulator of IL2).
GN TCF8 OR AREB6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=94186507; PubMed=8138542;
RX Watanabe Y., Kawakami K., Hirayama Y., Nagano K.;
RA "Transcription factors positively and negatively regulating the Na,K-
RT ATPase alpha 1 subunit gene";
RL J. Biochem. 114:849-855(1993).
RN [2]
RN SEQUENCE FROM N.A.
RP Bachman N.J., Scarpulla R.C.;
RA "A human zinc finger homeodomain protein homologous to the chicken
RT delta-crystallin enhancer binding protein, delta EPI";
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 390-1124 FROM N.A.

RX MEDLINE=92108424; PubMed=1840704;
RA Williams T.M., Moolten D., Burlin J., Romano J., Bhaerman R.,
RA Godillot A., Mellon M., Rauscher F.J. III, Kant J.A.;
RT "Identification of a zinc finger protein that inhibits IL-2 gene
RT expression";
RL Science 254:1791-1794(1991).
CC -1- FUNCTION: INHIBITS INTERLEUKIN-2 (IL-2) GENE EXPRESSION. MAY BE
CC RESPONSIBLE FOR TRANSCRIPTIONAL REPRESSION OF THE IL-2 GENE.
CC ENHANCES OR REPRESSES THE PROMOTER ACTIVITY OF THE ATRIAL GENE
CC DEPENDING ON THE QUANTITY OF CDNA AND ON THE CELL TYPE.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART AND SKELETAL MUSCLE, BUT
CC NOT IN LIVER, SPLEEN, OR PANCREAS.
CC -1- SIMILARITY: BELONGS TO DELTA-EF1/ZFP-1 FAMILY OF TWO-HANDED ZINC
CC FINGER/HOMEODOMAIN PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D15050; BAA03646.1; -;
CC EMBL; U12170; AAA20602.1; -;
CC EMBL; M81699; -; NOT_ANNOTATED_CDS.
CC PUR; JX0293; JX0293.
CC TRANSFAC; T00625; -;
CC GENE; HGNC:11642; TCF8.
CC MIM; 189909; -;
CC GO; GO:0003713; F:transcription co-activator activity; TAS.
CC GO; GO:0003714; F:transcription co-repressor activity; TAS.
CC GO; GO:0003700; F:transcription factor activity; TAS.
CC GO; GO:0008270; F:zinc ion binding; TAS.
CC GO; GO:0008283; P:cell proliferation; TAS.
CC GO; GO:0006955; P:immune response; TAS.
CC GO; GO:0001122; P:negative regulation of transcription from P...; TAS.
CC InterPro; IPR001356; Homeobox.
CC InterPro; IPR007087; Znf_C2H2.
CC Pfam; PF00096; zf-C2H2; 7.
CC ProDom; PD000010; Homeobox; 1.
CC ProDom; PD000003; Znf_C2H2; 2.
CC SMART; SM00389; Hox; 1.
CC SMART; SM00355; Znf_C2H2; 7.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
CC PROSITE; PS00157; ZINC_FINGER_C2H2_2; 6.
CC KW Transcription regulation; DNA-binding; Nuclear protein; Zinc-finger;
CC Homeobox; Repressor; Activator; Metal-binding; Repeat.
CC ZN_FING 170 193 C2H2-TYPE.
CC FT ZN_FING 200 222 C2H2-TYPE.
CC FT ZN_FING 240 262 C2H2-TYPE.
CC FT ZN_FING 268 292 C2H2-TYPE (ATYPICAL).
CC FT DNA_BIND 581 640 HOMEBOX-LIKE.
CC FT ZN_FING 904 926 C2H2-TYPE.
CC FT ZN_FING 932 954 C2H2-TYPE.
CC FT ZN_FING 960 981 C2H2-TYPE (ATYPICAL).
CC FT DOMAIN 989 1124 GLU-RICH (ACIDIC).
CC FT CONFLICT 420 420 V -> I (IN REF. 2).
CC FT CONFLICT 609 609 E -> Q (IN REF. 3).
CC FT CONFLICT 654 654 I -> T (IN REF. 2).
CC FT CONFLICT 672 672 D -> H (IN REF. 3).
CC FT CONFLICT 681 681 L -> S (IN REF. 3).
CC SQ SEQUENCE 1124 AA; 124073 MW; 0A2714CC37C848D1 CRC64;
Query Match 37.5%; Score 76.5; DB 1; Length 1124;
Best Local Similarity 40.9%; Pred. No. 11;
Matches 18; Conservative 7; Mismatches 8; Indels 11; Gaps 1;
QY 8 SSXDKENISKENDVDLDEK-----EEEAETEELEEK 40
DB 1036 SEKEEEEDKEEELQEKCEKPGQDEEEEEEVEEVEEVEE 1079

RESULT 12
RT ATRX_HUMAN STANDARD; PRT; 2492 AA.
RL ATRX_HUMAN
AC P46100; P51068; Q15886; Q9H021; Q9NTS3;
DT 01-NOV-1995 (Rel. 32, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-WAR-2004 (Rel. 43, Last annotation update)
DE Transcriptional regulator ATRX (X-linked helicase II) (X-linked
DE nuclear protein) (XNP) (Znf-HX).
GN ATRX OR RAD54L OR XH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5), VARIANTS SER-596;
RP GLU-740 AND SER-1860, AND VARIANTS ATR-X.
RX MEDLINE=97123494; PubMed=8968741;
RA Picketts D.J., Higgs D.R., Bachoo S., Blake D.J., Quarrell O.W.J.,
RA Gibbons R.J.;
RT "ATRX encodes a novel member of the SNF2 family of proteins: mutations
RT point to a common mechanism underlying the ATR-X syndrome.";
RL Hum. Mol. Genet. 5:1899-1907(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).
RX MEDLINE=97386582; PubMed=9244431;
RA Villard L., Lossi A.-M., Cardoso C., Proud V., Chiaroni P.,
RA Collea L., Schwartz C., Fontes M.;
RT "Determination of the genomic structure of the XNP/ATRX gene encoding
RT a potential zinc finger helicase.";
RL Genomics 43:149-155(1997).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANTS SER-596 AND GLU-740.
RX MEDLINE=22763540; PubMed=1277533;
RA Kitano T., Schwarz C., Nickel B., Paabo S.;
RT "Gene diversity patterns at 10 X-chromosomal loci in humans and
RT chimpanzees.";
RL Mol. Biol. Evol. 20:1281-1289(2003).
RN [4]
RP SEQUENCE OF 860-2492 FROM N.A.
RX MEDLINE=95179111; PubMed=7874112;
RA Stayton C.L., Dabovic B., Guliasso M., Gecz J., Broccoli V.,
RA Giovanazzi S., Bosolasco M., Monaco L., Raetan S., Boncinelli E.,
RA Bianchi M.E., Consalez G.G.;
RT "Cloning and characterization of a new human Xq13 gene, encoding a
RT putative helicase.";
RL Hum. Mol. Genet. 3:1957-1964(1994).
RN [5]
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=94214473; PubMed=8182050;
RA Gecz J., Pollard H., Consalez G., Villard L., Stayton C.L.,
RA Millassseau P., Khrestchatskiy M., Fontes M.;
RT "Cloning and expression of the murine homologue of a putative human
RT X-linked nuclear protein gene closely linked to PGK1 in Xq13.3.";
RL Hum. Mol. Genet. 3:39-44(1994).
RN [6]
RP SEQUENCE OF 2401-2492 FROM N.A., AND VARIANTS ATR-X.
RX MEDLINE=95211835; PubMed=7697714;
RA Gibbons R.J., Picketts D.J., Villard L., Higgs D.R.;
RT "Mutations in a putative global transcriptional regulator cause X-
RT linked mental retardation with alpha-thalassemia (ATR-X syndrome).";
RL Cell 80:837-845(1995).
RN [7]
RP SEQUENCE OF 1375-2492 FROM N.A.
RA Pearce A., Chapman J.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [8]
RP EZH2 BINDING.
RX MEDLINE=98167853; PubMed=9499421;
RA Cardoso C., Timsit S., Villard L., Khrestchatskiy M., Fontes M.,
RA Collea L.;
RT "Specific interaction between the XNP/ATRX gene product and the SET

RT domain of the human EZH2 protein.";
RL Hum. Mol. Genet. 7:679-684(1998).
RN [9]
RP SUBCELLULAR LOCATION, AND ASSOCIATION WITH PERICENTROMERIC
RP HETEROCHROMATIN.
RX MEDLINE=20040663; PubMed=10570185;
RA McDowell T.L., Gibbons R.J., Sutherland H., O'Rourke D.M.,
RA Rickmore W.A., Pombo A., Turley H., Gatter K., Picketts D.J.,
RA Buckle V.J., Chapman L., Rhodes D., Higgs D.R.;
RT "Localization of a putative transcriptional regulator (ATRX) at
RT pericentromeric heterochromatin and the short arms of acrocentric
RT chromosomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:13983-13988(1999).
RN [10]
RP DISEASE.
RX MEDLINE=20213147; PubMed=10751095;
RA Villard L., Fontes M., Ades L.C., Gecz J.;
RT "Identification of a mutation in the XNP/ATR-X gene in a family
RT reported as Smith-Fineman-Myers syndrome.";
RL Am. J. Med. Genet. 91:83-85(2000).
RN [11]
RP DISEASE.
RX MEDLINE=22804448; PubMed=12858175;
RA Gibbons R.J., Pellagatti A., Garrick D., Wood W.G., Malik N.,
RA Ayub H., Langford C., Boulwood J., Wainscoat J.S., Higgs D.R.;
RT "Identification of acquired somatic mutations in the gene encoding
RT chromatin-remodeling factor ATRX in the alpha-thalassemia
RT myelodysplasia syndrome (ATMDS).";
RL Nat. Genet. 34:446-449(2003).
RN [12]
RP VARIANT ATR-X SER-1713.
RX MEDLINE=97196774; PubMed=9043863;
RA Villard L., Lacombe D., Fontes M.;
RT "A point mutation in the XNP gene, associated with an ATR-X phenotype
RT without alpha-thalassemia.";
RL Eur. J. Hum. Genet. 4:316-320(1996).
RN [13]
RP VARIANT JM GLN-2131.
RX MEDLINE=96224392; PubMed=8630485;
RA Villard L., Gecz J., Mattel J.-F., Fontes M., Saugier-Verber P.,
RA Munnich A., Lyonnet S.;
RT "XNP mutation in a large family with Juberg-Marsidi syndrome.";
RL Nat. Genet. 12:359-360(1996).
RN [14]
RP VARIANTS ATR-X.
RX MEDLINE=97467722; PubMed=9326931;
RA Gibbons R.J., Bachoo S., Picketts D.J., Aftimos S., Asenbauer B.,
RA Bergoffen J., Berry S.A., Dahl N., Fryer A., Keppler K., Kurosawa K.,
RA Levin M.L., Masuno M., Neri G., Pierpont M.E., Slaney S.F.,
RA Higgs D.R.;
RT "Mutations in transcriptional regulator ATRX establish the functional
RT significance of a PHD-like domain.";
RL Nat. Genet. 17:146-148(1997).
RN [15]
RP VARIANT ATR-X LEU-246.
RX MEDLINE=20123062; PubMed=10660327;
RA Fichera M., Romano C., Castiglia L., Failla P., Ruberto C., Amata S.,
RA Greco D., Cardoso C., Fontes M., Ragusa A.;
RT "New mutations in XNP/ATR-X gene: a further contribution to
RT genotype/phenotype relationship in ATR/X syndrome.";
RL Hum. Mutat. 12:214-214(1998).
RN [16]
RP VARIANT SHS LYS-1742.
RX MEDLINE=99347960; PubMed=10417298;
RA Lossi A.-M., Millan J.M., Villard L., Orellana C., Cardoso C.,
RA Prieto F., Fontes M., Martinez F.;
RT "Mutation of the XNP/ATR-X gene in a family with severe mental
RT retardation, spastic paraplegia and skewed pattern of X inactivation:
RT demonstration that the mutation is involved in the inactivation
RT bias.";
RL Am. J. Hum. Genet. 65:558-562(1999).
RN [17]
RP VARIANT CWS THR-2050.

```

RX MEDLINE=99326063; PubMed=10398237;
RA Abidi F., Schwartz C.E., Carpenter N.J., Villard L., Fontes M.,
RA Curtis M.;
RT "Carpenter-Waziri syndrome results from a mutation in XNP.";
RL Am. J. Med. Genet. 85:249-251(1999).
RN [18]
RP VARIANTS ATR-X GLU-175; 178-VAL--LYS-198 DEL; SER-190; PRO-219;
RP LEU-246 AND CYS-249.
RX MEDLINE=99219535; PubMed=10204841;
RA Villard L., Bonino M.-C., Abidi F., Ragusa A., Beloungue J.,
RA Loessi A.-M., Seaver L., Bonnefont J.-P., Romano C., Fichera M.,
RA Lacombe D., Hanauer A., Philip N., Schwartz C.E., Fontes M.;
RT "Evaluation of a mutation screening strategy for sporadic cases of
RT ATR-X syndrome.";
RL J. Med. Genet. 36:183-186(1999).
RN [19]
RP VARIANTS ATR-X SER-179; LEU-190; ILE-194; CYS-246; PHE-1552; SER-1645
RP AND CYS-1847.
RX MEDLINE=20451413; PubMed=10995512;
RA Wada T., Kubota T., Fukushima Y., Saitoh S.;
RT "Molecular genetic study of Japanese patients with X-linked alpha-
RT thalassemia/mental retardation syndrome (ATR-X).";
RL Am. J. Med. Genet. 94:242-248(2000).
RN [20]
RP VARIANT ATR-X MET-1621.
RX MEDLINE=22108790; PubMed=12116232;
RA Yntema H.G., Poppelars F.A., Derksen E., Oudakker A.R.,
RA van Roosmalen T., Jacobs A., Obbema H., Brunner H.G., Hamel B.C.J.,
RA van Bokhoven H.;
RT "Expanding phenotype of XNP mutations: mild to moderate mental
RT retardation.";
RL Am. J. Med. Genet. 110:243-247(2002).
CC -1- FUNCTION: Could be a global transcriptional regulator. Modifies
CC gene expression by affecting chromatin. May be involved in brain
CC development and facial morphogenesis.
CC -1- SUBUNIT: Probably binds EZH2. Binds annexin V in a calcium and
CC phosphatidylcholine/phosphatidylserine-dependent manner (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear. Associated with pericentromeric
CC heterochromatin during interphase and mitosis, probably by
CC interacting with HPI.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Names=4;
CC IsoId=P46100-1; Sequence=Displayed;
CC Names=1;
CC IsoId=P46100-2; Sequence=VSP_000575;
CC Names=2;
CC IsoId=P46100-3; Sequence=VSP_000574;
CC Names=3;
CC IsoId=P46100-4; Sequence=VSP_000576;
CC Names=5;
CC IsoId=P46100-5; Sequence=VSP_000574, VSP_000576;
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- DISEASE: Defects in ATRX are the cause of X-linked alpha-
CC thalassemia/mental retardation syndrome (ATR-X) [MIM:301040]. ATR-
CC X is an X-linked disorder comprising severe psychomotor
CC
Query Match 37.3%; Score 76; DB 1; Length 2492;
Best Local Similarity 52.9%; Pred. No. 27;
Matches 18; Conservative 5; Mismatches 7; Indels 4; Gaps 1;
QY 8 SSKDKENISKENDVDLDEKEEAEETEELEEK 41
Db 1435 SSENKSNSEEE-----EKEEEEEEEEEEEEE 1464
RESULT 13
ATR_X_PANTR STANDARD; PRT; 2492 AA.
AC QYQM4;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)

```

```

DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Transcriptional regulator ATRX (X-linked helicase II) (X-linked
DE nuclear protein) (XNP).
GN ATRX.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22763540; PubMed=12777533;
RA Kitano T., Schwarz C., Nickel B., Paabo S.;
RT "Gene diversity patterns at 10 X-chromosomal loci in humans and
RT chimpanzees.";
RL Mol. Biol. Evol. 20:1281-1289(2003).
CC -1- FUNCTION: Could be a global transcriptional regulator. Modifies
CC gene expression by affecting chromatin. May be involved in brain
CC development and facial morphogenesis.
CC -1- SUBUNIT: Probably binds EZH2. Binds annexin V in a calcium and
CC phosphatidylcholine/phosphatidylserine-dependent manner (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear. Associated with pericentromeric
CC heterochromatin during interphase and mitosis, probably by
CC interacting with HPI.
CC -1- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.
CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB102642; BAC91111.1;
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR000330; SNF2_N.
DR Pfam; PF00271; Helicase_C_1.
DR Pfam; PF00176; SNF2_N_1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR DNA repair; Nuclear protein; DNA-binding; Helicase; ATP-binding;
KW Zinc-finger.
FT ZN_FING 220 268 PHD-TYPE.
FT NP_BIND 1594 1601 ATP (POTENTIAL).
FT SITE 1719 1722 DEGR_BOX.
FT DOMAIN 745 750 POLY-SER.
FT DOMAIN 1151 1156 POLY-SER.
FT DOMAIN 1166 1169 POLY-LYS.
FT DOMAIN 1202 1206 POLY-SER.
FT DOMAIN 1259 1266 POLY-ASP.
FT DOMAIN 1443 1466 POLY-GLU.
FT DOMAIN 1499 1502 POLY-GLU.
FT DOMAIN 1929 1939 POLY-LYS.
FT DOMAIN 1941 1948 POLY-SER.
FT DOMAIN 2221 2225 POLY-LYS.
FT DOMAIN 2262 2265 POLY-GLU.
FT DOMAIN 2420 2425 POLY-GLN.
SQ SEQUENCE 2492 AA; 282567 MW; 56978AC7D37400DC CRC64;
Query Match 37.3%; Score 76; DB 1; Length 2492;
Best Local Similarity 52.9%; Pred. No. 27;
Matches 18; Conservative 5; Mismatches 7; Indels 4; Gaps 1;
QY 8 SSKDKENISKENDVDLDEKEEAEETEELEEK 41
Db 1435 SSENKSNSEEE-----EKEEEEEEEEEEEEE 1464
RESULT 14
ATR_X_PONPY

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:09:13 ; Search time 88.438 Seconds
(without alignments)
148.878 Million cell updates/sec

Title: US-10-774-602-14
Perfect score: 204
Sequence: 1.MLSHLVSSKDKENISKEND.....VLDEKEAEETEELEEK 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 32113274 residues

Total number of hits satisfying chosen parameters: 1342398

Maximum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	204	100.0	41	14	US-10-294-770-14
2	204	100.0	41	16	US-10-774-602-14
3	101	49.5	28	14	US-10-294-770-14
4	101	49.5	28	16	US-10-774-602-14
5	97	47.5	28	14	US-10-294-770-14
6	97	47.5	28	16	US-10-294-770-14
7	97	47.5	28	14	US-10-294-770-14
8	97	47.5	28	16	US-10-294-770-14
9	97	47.5	64	14	US-10-294-770-14
10	97	47.5	64	16	US-10-294-770-14
11	87	42.6	1077	15	US-10-099-322-110
12	87	42.6	1077	15	US-10-044-564-110
13	83	40.7	197	13	US-10-101-487-51
14	83	40.7	197	13	US-10-101-487-114
15	80	39.2	161	16	US-10-437-963-108765

16	79	38.7	114	9	US-09-864-761-36168	Sequence 36168, A
17	79	38.7	181	13	US-10-101-487-45	Sequence 45, Appl
18	79	38.7	714	9	US-09-978-242-3	Sequence 3, Appl
19	79	38.7	741	12	US-10-282-122A-65999	Sequence 65999, A
20	78	38.2	89	9	US-09-864-761-44120	Sequence 44120, A
21	78	38.2	117	9	US-09-864-761-45592	Sequence 45592, A
22	78	38.2	225	14	US-10-029-386-32045	Sequence 32045, A
23	78	38.2	406	16	US-10-437-963-115969	Sequence 115969, A
24	77	37.7	186	13	US-10-101-487-44	Sequence 44, Appl
25	77	37.7	198	13	US-10-101-487-42	Sequence 42, Appl
26	77	37.7	200	13	US-10-101-487-53	Sequence 53, Appl
27	76.5	37.5	97	14	US-10-029-386-33853	Sequence 33853, A
28	76.5	37.5	1104	15	US-10-104-047-2506	Sequence 2506, Ap
29	76.5	37.5	1125	9	US-09-974-298-114	Sequence 114, App
30	76.5	37.5	1125	12	US-10-087-192-234	Sequence 234, App
31	76.5	37.5	1154	14	US-10-232-561-10	Sequence 10, Appl
32	76	37.3	179	13	US-10-101-487-107	Sequence 107, App
33	76	37.3	983	12	US-10-112-944-435	Sequence 435, App
34	76	37.3	1021	16	US-10-408-765A-313	Sequence 313, App
35	76	37.3	1641	16	US-10-408-765A-814	Sequence 814, App
36	76	37.3	1641	16	US-10-408-765A-815	Sequence 815, App
37	76	37.3	2375	16	US-10-408-765A-277	Sequence 277, App
38	76	37.3	2492	16	US-10-697-526-2	Sequence 2, Appl
39	75.5	37.0	319	12	US-10-425-114-44044	Sequence 44044, A
40	75.5	37.0	1257	16	US-10-408-765A-1486	Sequence 1486, Ap
41	75	36.8	59	14	US-10-029-386-32735	Sequence 32735, A
42	75	36.8	214	9	US-09-214-881A-5	Sequence 5, Appl
43	75	36.8	1081	16	US-10-408-765A-298	Sequence 298, App
44	75	36.8	1251	12	US-10-112-944-881	Sequence 881, App
45	75	36.8	1270	16	US-10-408-765A-2253	Sequence 2253, Ap

ALIGNMENTS

RESULT 1
US-10-294-770-14
; Sequence 14, Application US/10294770
; Publication No. US20030161840A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759USOCIP
; CURRENT APPLICATION NUMBER: US/10/294, 770
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 14
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-294-770-14

Query Match 100.0%; Score 204; DB 14; Length 41;
Best Local Similarity 100.0%; Pred. No. 5.9e-13;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLVSSKDKENISKENDVLDEKEAEETEELEEK 41
|||||
Db 1 MLSHLVSSKDKENISKENDVLDEKEAEETEELEEK 41

RESULT 2
US-10-774-602-14
; Sequence 14, Application US/10774602

; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791USODIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-774-602-14

Query Match 100.0%; Score 204; DB 16; Length 41;
Best Local Similarity 100.0%; Pred. No. 5.9e-13;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLSHLYVSSKDKENISKENDVLDKEEAEAEETEEELSEK 41
DB 1 MLSHLYVSSKDKENISKENDVLDKEEAEAEETEEELSEK 41

RESULT 3
US-10-294-770-13
; Sequence 13, Application US/10294770
; Publication No. US20030161840A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759USOCIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-294-770-13

Query Match 49.5%; Score 101; DB 14; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLSHLYVSSKDKENISKEND 20
DB 9 MLSHLYVSSKDKENISKEND 28

RESULT 4
US-10-774-602-13
; Sequence 13, Application US/10774602
; Publication No. US20040141987A1

; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791USODIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-774-602-13

Query Match 49.5%; Score 101; DB 16; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLSHLYVSSKDKENISKEND 20
DB 9 MLSHLYVSSKDKENISKEND 28

RESULT 5
US-10-294-770-4
; Sequence 4, Application US/10294770
; Publication No. US20030161840A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759USOCIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-294-770-4

Query Match 47.5%; Score 97; DB 14; Length 28;
Best Local Similarity 95.0%; Pred. No. 0.0078;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLSHLYVSSKDKENISKEND 20
DB 9 MLSHLYVSSKDKENISKENE 28

RESULT 6
US-10-238-741-4
; Sequence 4, Application US/10238741
; Publication No. US20040096466A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAQ

OEUVRAY, CLAUDE
TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: DRUILHE, PIERRE
FILING DATE: 19-JUL-1999
APPLICATION NUMBER: US/09/356,497
FILING DATE: 08-AUG-1995
APPLICATION NUMBER: US/08/416,711
FILING DATE: 18-OCT-1993
APPLICATION NUMBER: PCT/FR93/01024
FILING DATE: 19-OCT-1992
APPLICATION NUMBER: FR 92/12488
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-085-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-238-741-4
Query Match 47.5%; Score 97; DB 16; Length 28;
Best Local Similarity 95.0%; Pred. No. 0.0078;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLSHLYVSSKDKENISKEND 20
Db 9 MLSHLYVSSKDKENISKENE 28
RESULT 7
US-10-774-602-4
Sequence 4, Application US/10774602
Publication No. US20040141987A1
GENERAL INFORMATION:
APPLICANT: DRUILHE, PIERRE
TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
FILE REFERENCE: 248791USODIV
CURRENT APPLICATION NUMBER: US/10/774,602
CURRENT FILING DATE: 2004-02-10
PRIOR APPLICATION NUMBER: US 09/356,947
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: US 10/238,741
PRIOR FILING DATE: 2002-09-11
PRIOR APPLICATION NUMBER: US 08/416,711
PRIOR FILING DATE: 1995-08-08
PRIOR APPLICATION NUMBER: PCT/FR93/01024

PRIOR FILING DATE: 1993-10-18
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patent in version 3.1
SEQ ID NO 4
LENGTH: 28
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-10-774-602-4
Query Match 47.5%; Score 97; DB 16; Length 28;
Best Local Similarity 95.0%; Pred. No. 0.0078;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLSHLYVSSKDKENISKEND 20
Db 9 MLSHLYVSSKDKENISKENE 28
RESULT 8
US-10-294-770-1
Sequence 1, Application US/10294770
Publication No. US20030161840A1
GENERAL INFORMATION:
APPLICANT: DRUILHE, PIERRE
TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
FILE REFERENCE: 230759USOCIP
CURRENT APPLICATION NUMBER: US/10/294,770
CURRENT FILING DATE: 2002-11-15
PRIOR APPLICATION NUMBER: US 09/356,947
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: US 08/416,711
PRIOR FILING DATE: 1995-08-08
PRIOR APPLICATION NUMBER: PCT/FR93/01024
PRIOR FILING DATE: 1993-10-18
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patent in version 3.1
SEQ ID NO 1
LENGTH: 64
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-10-294-770-1
Query Match 47.5%; Score 97; DB 14; Length 64;
Best Local Similarity 95.0%; Pred. No. 0.018;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLSHLYVSSKDKENISKEND 20
Db 45 MLSHLYVSSKDKENISKENE 64
RESULT 9
US-10-238-741-1
Sequence 1, Application US/10238741
Publication No. US20040096466A1
GENERAL INFORMATION:
APPLICANT: DRUILHE, PIERRE
BOUHAROUN-TAYOUN, HASNAQ
OEUVRAY, CLAUDE
TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: DRUILHE, PIERRE
FILING DATE: 19-JUL-1999
APPLICATION NUMBER: US/09/356,497
FILING DATE: 08-AUG-1995
APPLICATION NUMBER: US/08/416,711
FILING DATE: 18-OCT-1993
APPLICATION NUMBER: PCT/FR93/01024
FILING DATE: 19-OCT-1992
APPLICATION NUMBER: FR 92/12488
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-085-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-238-741-4
Query Match 47.5%; Score 97; DB 16; Length 28;
Best Local Similarity 95.0%; Pred. No. 0.0078;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLSHLYVSSKDKENISKEND 20
Db 9 MLSHLYVSSKDKENISKENE 28
RESULT 7
US-10-774-602-4
Sequence 4, Application US/10774602
Publication No. US20040141987A1
GENERAL INFORMATION:
APPLICANT: DRUILHE, PIERRE
TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
FILE REFERENCE: 248791USODIV
CURRENT APPLICATION NUMBER: US/10/774,602
CURRENT FILING DATE: 2004-02-10
PRIOR APPLICATION NUMBER: US 09/356,947
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: US 10/238,741
PRIOR FILING DATE: 2002-09-11
PRIOR APPLICATION NUMBER: US 08/416,711
PRIOR FILING DATE: 1995-08-08
PRIOR APPLICATION NUMBER: PCT/FR93/01024

/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA: US/10/238,741
/ APPLICATION NUMBER: US/10/238,741
/ FILING DATE: 09-Nov-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/09/356,497
/ FILING DATE: 19-Jul-1999
/ APPLICATION NUMBER: US/08/416,711
/ FILING DATE: 08-Aug-1995
/ APPLICATION NUMBER: PCT/FR93/01024
/ FILING DATE: 18-Oct-1993
/ APPLICATION NUMBER: FR 92/12488
/ FILING DATE: 19-Oct-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: OBLON, NORMAN F.
/ REGISTRATION NUMBER: 24,618
/ REFERENCE/DOCKET NUMBER: 660-085-0 PCT
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 703-413-3000
/ TELEFAX: 703-413-2220
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 64 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-238-741-1

Query Match 47.5%; Score 97; DB 16; Length 64;
Best Local Similarity 95.0%; Pred. No. 0.018;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
DB 45 MLSHLYVSSKDKENISKENE 64

RESULT 10
US-10-774-602-1
/ Sequence 1, Application US/10774602
/ Publication No. US20040141987A1
/ GENERAL INFORMATION:
/ APPLICANT: DRIULHE, PIERRE
/ TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
/ FILE REFERENCE: 248791US01V
/ CURRENT APPLICATION NUMBER: US/10/774,602
/ CURRENT FILING DATE: 2004-02-10
/ PRIOR APPLICATION NUMBER: US 09/356,947
/ PRIOR FILING DATE: 1999-07-19
/ PRIOR APPLICATION NUMBER: US 10/238,741
/ PRIOR FILING DATE: 2002-09-11
/ PRIOR APPLICATION NUMBER: US 08/416,711
/ PRIOR FILING DATE: 1995-08-08
/ PRIOR APPLICATION NUMBER: PCT/FR93/01024
/ PRIOR FILING DATE: 1993-10-18
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 64
/ TYPE: PRT
/ ORGANISM: Plasmodium falciparum
US-10-774-602-1

Query Match 47.5%; Score 97; DB 16; Length 64;
Best Local Similarity 95.0%; Pred. No. 0.018;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
DB 45 MLSHLYVSSKDKENISKENE 64

DB 45 MLSHLYVSSKDKENISKENE 64
RESULT 11
US-10-099-322-110
/ Sequence 110, Application US/10099322
/ Publication No. US20030215449A1
/ GENERAL INFORMATION:
/ APPLICANT: Mezes et al.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-240CIP
/ CURRENT APPLICATION NUMBER: US/10/099,322
/ CURRENT FILING DATE: 2002-09-11
/ PRIOR APPLICATION NUMBER: 60/261,014
/ PRIOR FILING DATE: 2001-01-11
/ PRIOR APPLICATION NUMBER: 60/261,018
/ PRIOR FILING DATE: 2001-01-11
/ PRIOR APPLICATION NUMBER: 60/318,410
/ PRIOR FILING DATE: 2001-09-10
/ PRIOR APPLICATION NUMBER: 60/261,013
/ PRIOR FILING DATE: 2001-01-11
/ PRIOR APPLICATION NUMBER: 60/261,026
/ PRIOR FILING DATE: 2001-01-11
/ PRIOR APPLICATION NUMBER: 60/261,029
/ PRIOR FILING DATE: 2001-01-11
/ PRIOR APPLICATION NUMBER: 60/313,170
/ PRIOR FILING DATE: 2001-08-17
/ PRIOR APPLICATION NUMBER: 10/044,564
/ PRIOR FILING DATE: 2002-01-11
/ NUMBER OF SEQ ID NOS: 324
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 110
/ LENGTH: 1077
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-10-099-322-110

Query Match 42.6%; Score 87; DB 15; Length 1077;
Best Local Similarity 55.9%; Pred. No. 3.1;
Matches 19; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 8 SSKDKENISKENDVDLDEKEEAETEEEEK 41
DB 970 ASKSDDESESENDDEDEEESESESEEEK 1003

RESULT 12
US-10-044-564-110
/ Sequence 110, Application US/10044564
/ Publication No. US20040018196A1
/ GENERAL INFORMATION:
/ APPLICANT: Mezes et al.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-240
/ CURRENT APPLICATION NUMBER: US/10/044,564
/ CURRENT FILING DATE: 2002-09-09
/ PRIOR APPLICATION NUMBER: 60/261,014
/ PRIOR FILING DATE: 2001-01-11
/ PRIOR APPLICATION NUMBER: 60/261,018
/ PRIOR FILING DATE: 2001-01-11
/ PRIOR APPLICATION NUMBER: 60/318,410
/ PRIOR FILING DATE: 2001-09-10
/ PRIOR APPLICATION NUMBER: 60/261,013
/ PRIOR FILING DATE: 2001-01-11
/ PRIOR APPLICATION NUMBER: 60/261,026
/ PRIOR FILING DATE: 2001-01-11
/ PRIOR APPLICATION NUMBER: 60/261,029
/ PRIOR FILING DATE: 2001-01-11
/ PRIOR APPLICATION NUMBER: 60/313,170
/ PRIOR FILING DATE: 2001-08-17
/ NUMBER OF SEQ ID NOS: 306
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 110

53



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:00:42 ; Search time 65.3967 Seconds
(without alignments)
177.141 Million cell updates/sec

Title: US-10-774-602-14
Perfect score: 204
Sequence: 1 MLSHLYVSKKENISKEND.....VLDEKKEAEETEEBLEEK 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	47.5	28	2 AAR54687	Aar54687 Plasmodiu
2	97	47.5	64	2 AAR54684	Aar54684 Plasmodiu
3	97	47.5	64	6 ABR41813	Abri41813 Merozoite
4	87	42.6	1077	4 ABB58720	Abb58720 Drosophil
5	83	40.7	31	6 ABR41855	Abri41855 Peptide u
6	83	40.7	89	4 ABG04390	Abg04390 Novel hum
7	83	40.7	106	4 ABG06955	Abg06955 Novel hum
8	83	40.7	197	6 ABP56385	Abp56385 Polyanion
9	83	40.7	379	4 ABG26552	Abg26552 Novel hum
10	83	40.7	788	4 ABG29229	Abg29229 Novel hum
11	82	40.2	31	6 ABR41854	Abri41854 Peptide u
12	82	40.2	57	4 ABG11240	Abg11240 Novel hum
13	82	40.2	62	4 ABG03970	Abg03970 Novel hum
14	82	40.2	295	4 ABG11244	Abg11244 Novel hum
15	82	40.2	412	3 AAG19532	Aag19532 Arabidops
16	82	40.2	414	3 AAG19531	Aag19531 Arabidops
17	82	40.2	456	3 AAG19530	Aag19530 Arabidops
18	82	40.2	3135	2 AAG57474	Aag57474 P. falcip
19	82	40.2	3135	3 AAB18223	Aab18223 Plasmodiu
20	81	39.7	81	3 AAG37760	Aag37760 Arabidops
21	81	39.7	81	3 AAG15127	Aag15127 Arabidops
22	81	39.7	81	3 AAG37763	Aag37763 Arabidops
23	81	39.7	105	3 AAG36862	Aag36862 Arabidops
24	81	39.7	105	3 AAG37759	Aag37759 Arabidops
25	81	39.7	105	3 AAG37762	Aag37762 Arabidops

26	81	39.7	105	3 AAG15126	Aag15126 Arabidops
27	81	39.7	120	3 AAG36861	Aag36861 Arabidops
28	81	39.7	120	3 AAG15125	Aag15125 Arabidops
29	81	39.7	120	3 AAG37761	Aag37761 Arabidops
30	81	39.7	120	3 AAG37758	Aag37758 Arabidops
31	81	39.7	382	4 ABG04351	Abg04351 Novel hum
32	80	39.2	31	6 ABR41853	Abri41853 Peptide u
33	80	39.2	65	4 ABG06896	Abg06896 Novel hum
34	80	39.2	85	4 ABG05351	Abg05351 Novel hum
35	80	39.2	93	4 ABG02968	Abg02968 Novel hum
36	80	39.2	93	4 ABG04398	Abg04398 Novel hum
37	80	39.2	93	4 ABG26512	Abg26512 Novel hum
38	80	39.2	98	4 ABG11735	Abg11735 Novel hum
39	80	39.2	109	4 ABG06515	Abg06515 Novel hum
40	80	39.2	109	4 ABG06958	Abg06958 Novel hum
41	80	39.2	116	4 ABG11262	Abg11262 Novel hum
42	80	39.2	226	4 ABG15564	Abg15564 Novel hum
43	80	39.2	234	4 ABG26550	Abg26550 Novel hum
44	80	39.2	1229	4 ABG24366	Abg24366 Novel hum
45	79	38.7	53	4 ABG06901	Abg06901 Novel hum

ALIGNMENTS

RESULT 1
AAR54687
ID AAR54687 standard; protein; 28 AA.

AC AAR54687;
XX
DT 25-MAR-2003 (revised)
DT 15-DEC-1994 (first entry)
XX
DE Plasmodium falciparum merozoite 48KD surface protein fragment IV.
XX
KW Merozoite 48KD surface antigen; immunogen; vaccine; malaria;
KW cytophilic antibody.
XX
OS Plasmodium falciparum.
XX
PN WO9409140-A1.
XX
PD 28-APR-1994.
XX
PF 18-OCT-1993; 93WO-FR001024.
XX
PR 19-OCT-1992; 92FR-00012488.
(INSP) INST PASTEUR.
Druilhe P, Bouharoun-Tayoun H, Oeuvery C;
WPI; 1994-151325/18.
N-PSDB; AAO64680.
New Plasmodium antigen generating cytophilic - mimicking protection
induced by long term exposure to the parasite, useful in vaccination
against and diagnosis of malaria, and prodn of antibodies for passive
immunisation.
Claim 4; Page 37; 52pp; French.
This sequence is a preferred fragment of the 48KD merozoite surface
antigen of P.falciparum which is useful as an immunogen to generate
antibodies (Abs). The Abs can induce a cytotoxic reaction against the
erythrocyte stage of the parasite. (Updated on 25-MAR-2003 to correct PN
field.)
XX
SQ Sequence 28 AA;

Query Match 47.5%; Score 97; DB 2; Length 28;
Best Local Similarity 95.0%; Pred. No. 0.00065;

Matches	19; Conservative	1; Mismatches	0; Indels	0; Gaps	0;
Qy	1	MLSHLYVSSKDKENISKEND 20			
Db	9	MLSHLYVSSKDKENISKENE 28			
RESULT 2					
AAAR54684					
ID	AAAR54684	standard; protein; 64 AA.			
XX	XX	AAAR54684;			
AC	XX	25-MAR-2003 (revised)			
DT	DT	15-DEC-1994 (first entry)			
DE	XX	Plasmodium falciparum merozoite 48kD surface protein fragment I.			
XX	XX	Merozoite 48kD surface antigen; immunogen; vaccine; malaria;			
XX	XX	cytophilic antibody.			
OS	XX	Plasmodium falciparum.			
XX	XX	WO9409140-A1.			
PN	XX	28-APR-1994.			
PD	XX	18-OCT-1993; 93WO-FR001024.			
XX	XX	19-OCT-1992; 92FR-00012488.			
PF	XX	(INSP) INST PASTEUR.			
PR	XX	Druilhe P, Bouharoun-Tayoun H, Oeuvray C;			
XX	XX	WPI; 1994-151325/18.			
XX	XX	N-PSDB; AAQ64677.			
XX	XX	New Plasmodium antigen generating cytophilic - mimicking protection			
PT	PT	induced by long term exposure to the parasite, useful in vaccination			
PT	PT	against and diagnosis of malaria, and prodn of antibodies for passive			
PT	PT	immunisation.			
XX	XX	Claim 3; Page 36; 52pp; French.			
XX	XX	This sequence is a preferred fragment of the 48kD merozoite surface			
XX	XX	antigen of P.falciparum which is useful as an immunogen to generate			
XX	XX	antibodies (Abs). The Abs can induce a cytotoxic reaction against the			
XX	XX	erythrocyte stage of the parasite. Specific subfragments of AAR54684 are			
XX	XX	separately claimed (see AAR54685-R54687). (Updated on 25-MAR-2003 to			
XX	XX	correct PN field.)			
XX	XX	Sequence 64 AA;			
Qy	1	MLSHLYVSSKDKENISKEND 20			
Db	45	MLSHLYVSSKDKENISKENE 64			
RESULT 3					
ABR41813					
ID	ABR41813	standard; protein; 64 AA.			
XX	XX	ABR41813;			
AC	XX	11-AUG-2003 (first entry)			
XX	XX	Merozoite surface antigen MSP-3 peptide (aa194-257).			
DE	XX				
XX	XX				

PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR N-PSDB; ABL02823.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 2952; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1077 AA;

Query Match 42.6%; Score 87; DB 4; Length 1077;
Best Local Similarity 55.9%; Pred. No. 0.56;
Matches 19; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 8 SSKDKENISKENDVLDKEKEEAEETEEBLEEK 41
DB 970 ASKSDSESENDDDDDEEESESEEEEEK 1003

RESULT 5
ABR41855
ID ABR41855 standard; peptide; 31 AA.
XX
AC ABR41855;
XX
XX
DT 11-AUG-2003 (first entry)
XX
XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
XX
XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
XX immunotherapy; vaccine; epitope.
XX
XX Plasmodium falciparum.
XX
XX Key Location/Qualifiers
FH Misc-difference 15..16
FT /note="linker residues"
XX
XX GB2378949-A.
XX
XX 26-FEB-2003.
XX
XX 16-AUG-2001; 2001GB-00020057.
XX
XX 16-AUG-2001; 2001GB-00020057.
XX
XX (DZIE/) DZIEGIEL M S H.
PA (LUND/) LUNDQUIST R.
PA (NIEL/) NIELSEN L K.
XX
XX Driegiel MSH, Lundquist R, Nielsen LK;
PI
XX

DR WPI; 2003-303108/30.
XX
PT Novel recombinant human antibody specific for merozoite surface protein-3
PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
PT a malarial disease.
XX
XX Example 3; Fig 10; 75pp; English.
XX
XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
CC 15 and 16 replaced by proprietary linker residues. The peptides were
CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
CC encoded by clones obtained from a phage display system based on antibody
CC variable region genes from the peripheral blood leucocytes of malaria
CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
CC terminal part of the MSP-3 antigen, with highest reactivity with the
CC middle part. RAM1 reacted only weakly with the peptides. The invention
CC provides recombinant human antibodies against MSP-3, which are useful for
CC the diagnosis and treatment of malaria. The recombinant human antibodies
CC include the complementarity determining regions from the light and heavy
CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
XX
SQ Sequence 31 AA;

Query Match 40.7%; Score 83; DB 6; Length 31;
Best Local Similarity 85.0%; Pred. No. 0.025;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
DB 12 MLSXXVYSSKDKENISKENE 31

RESULT 6
ABG04390
ID ABG04390 standard; protein; 89 AA.
XX
AC ABG04390;
XX
XX 13-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #4381.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS68577.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 34749; 103pp; English.
XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX
 SQ Sequence 89 AA;

Query Match 40.7%; Score 83; DB 4; Length 89;
 Best Local Similarity 51.4%; Pred. No. 0.086;
 Matches 18; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 7 VSKKKNISKENDVDLDEKEEAETEELEEK 41
 ||||| : : : : : ||||| : : : : :
 Db 41 VSKKKEKKKEEEEEEEEEEEEEEEEEEE 75

RESULT 7
 ABG06955
 ID ABG06955 standard; protein; 106 AA.

AC ABG06955;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #6946.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

QY WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS71142.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 20; SEQ ID NO 37314; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX
 SQ Sequence 106 AA;

Query Match 40.7%; Score 83; DB 4; Length 106;
 Best Local Similarity 48.6%; Pred. No. 0.11;
 Matches 17; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 7 VSKKKNISKENDVDLDEKEEAETEELEEK 41
 ||||| : : : : : ||||| : : : : :
 Db 41 VSEKEEEDSEEEEEEEEEEEEEEEEEEE 75

RESULT 8
 ABP56385
 ID ABP56385 standard; peptide; 197 AA.

AC ABP56385;
 XX
 DT 11-MAR-2003 (first entry)
 XX
 DE Polyanionic fusion protein product #5.
 KW Polyanionic polymer; bioactivity; water solubility.

XX Synthetic.

XX WO200277036-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US008614.

XX 21-MAR-2001; 2001US-0277705P.

XX (LEUN/) LEUNG D W.

XX Leung DW, Bergman PA, Lofquist A, Pietz GE, Tompkins CK;

XX Waggoner DW;

XX WPI; 2003-058367/05.

XX Producing monodispersed preparation of polyanionic polymer for therapy,
 PT by expressing vector comprising ligation product of oligonucleotides
 PT encoding glutamate/aspartate residues in host cell and isolating the
 PT product.

XX Example 7; Page 46; 74pp; English.

XX The present invention describes a method (M) for producing a
 CC monodispersed preparation of a polyanionic polymer (PP) larger than 10
 CC kD. (M) involves inserting into an expression vector (EV) a ligation
 CC product formed by ligating together oligonucleotides that encode
 CC glutamate/aspartate residues, expressing EV in a host cell, and isolating
 CC the protein product (P) of EV, where (P) is PP and at least 80% of PP is

CC approximately of the same molecular weight. Also described: (1) a
CC recombinant fusion protein (I) comprising a polyanionic polypeptide and
CC another polypeptide at either one end or at both ends of it; (2) a
CC polyanionic polymer (II) conjugate comprising a polyanionic polymer and
CC leukine, where the polyanionic polymer is polyglutamic acid or
CC polyaspartic acid; (3) a vector (III) comprising a cassette which
CC comprises a nucleotide sequence encoding a polyanionic polymer and at
CC least one other nucleotide sequence, where the polyanionic polymer is
CC polyglutamic acid or polyaspartic acid; (4) production of (I); (5) a cell
CC (IV) comprising (III) or a vector that comprises a nucleotide sequence
CC that encodes a polyanionic polymer that is larger than 10 kDa; and (6) a
CC recombinantly-produced polyanionic polymer (V) that is of any molecular
CC weight or is larger than 10 kDa, and is conjugated to another protein. (I)
CC is useful for treating a disease or ailment in an individual by
CC administering (I) to the individual. (I) is also useful for delivering an
CC effective amount of a pharmaceutically active agent, a therapeutic
CC protein or a drug to a patient in need of it, or for diagnostic and
CC testing or research purposes. AB222045 to AB222131 and ABP6374 to
CC ABP56400 represent sequences used in the exemplification of the present
CC invention

XX Sequence 197 AA;

Query Match 40.7%; Score 83; DB 6; Length 197;

Best Local Similarity 41.5%; Pred. No. 0.22;

Matches 17; Conservative. 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVLDKEKEEAETEEELEEK 41

DB 10 MPENLYFGQEEEEEEEEEEEEEEEEEEEEEEEEEEEE 50

RESULT 9

ABG26552
ID ABG26552 standard; protein; 379 AA.

XX AC ABG26552;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #26543.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS90739.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.

XX PS Claim 20; SEQ ID NO 56911; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX in diagnostics as expressed sequence tags for identifying expressed

CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX CC
XX SQ Sequence 379 AA;

Query Match 40.7%; Score 83; DB 4; Length 379;

Best Local Similarity 51.4%; Pred. No. 0.46;

Matches 18; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 7 VSSKDKENISKENDVLDKEKEEAETEEELEEK 41

DB 13 VSKKXKKKEEEEEEEEEEEEEEEEEEEEEEE 47

RESULT 10

ABG29229
ID ABG29229 standard; protein; 788 AA.

XX AC ABG29229;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #29220.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS93416.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.

XX PS Claim 20; SEQ ID NO 59588; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed

genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have application in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Query Match 40.7%; Score 83; DB 4; Length 788;
Best Local Similarity 51.4%; Pred. No. 1.1;
Matches 18; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Sequence 788 AA;
1 MLSHLYVSSKDKENISKEN 19
13 MLXLYVSSKDKENISKEN 31

RESULT 11
ABR41854
ID ABR41854 standard; peptide; 31 AA.
XX ABR41854;
DT 11-AUG-2003 (first entry)
XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
DE Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
KW immunotherapy; vaccine; epitope.
XX Plasmodium falciparum.
XX Key Location/Qualifiers
FH Misc-difference 15..16 /notes="linker residues"
FT GB2378949-A.
XX
PD 26-FEB-2003.
XX
PF 16-AUG-2001; 2001GB-00020057.
XX
PR 16-AUG-2001; 2001GB-00020057.
XX (DZIE/) DZIEGIEL M S H.
PA (LUND/) LUNDQUIST R.
PA (NIEL/) NIELSEN L K.
XX
PI Dzigiel MSH, Lundquist R, Nielsen LK;
XX WPI; 2003-303108/30.
DR
XX Novel recombinant human antibody specific for merozoite surface protein-3 antigen of Plasmodium falciparum, useful for treatment and prophylaxis of a malarial disease.
XX
XX Example 3; Fig 10; 75pp; English.
PS
XX The present sequence is one of a set of 34 peptides (see ABR41822-55) used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues

15 and 16 replaced by proprietary linker residues. The peptides were examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are encoded by clones obtained from a phage display system based on antibody variable region genes from the peripheral blood leukocytes of malaria immune human individuals. RAM2 reacted with the N-terminal, middle and C-terminal part of the MSP-3 antigen, with highest reactivity with the middle part. RAM1 reacted only weakly with the peptides. The invention provides recombinant human antibodies against MSP-3, which are useful for the diagnosis and treatment of malaria. The recombinant human antibodies include the complementarity determining regions from the light and heavy chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

Sequence 31 AA;
Query Match 40.2%; Score 82; DB 6; Length 31;
Best Local Similarity 89.5%; Pred. No. 0.033;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEN 19
DB 13 MLXLYVSSKDKENISKEN 31

RESULT 12
ABG11240
ID ABG11240 standard; protein; 57 AA.
XX AC ABG11240;
DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #11231.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
OS Homo sapiens.
XX WO200175067-A2.
PN 11-OCT-2001.
XX
PD 30-MAR-2001; 2001WO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
DR N-PSDB; AAS75427.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.
PS Claim 20; SEQ ID NO 41599; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 57 AA;
SQ

```

SQ      Sequence 57 AA;
      Query Match      40.2%;      Score 82;      DB 4;      Length 57;
      Best Local Similarity 42.9%;      Pred. NO. 0.066;
      Matches 18;      Conservative 9;      Mismatches 11;      Indels

QY      4      HLY-----VSKOKENISKENDVDLDSKEEAETEEEEK 41
Db      2      HLYSSNEVHAKKEEEEEEEEEEEEEEEEEEEEEEEEEEEEE 43

```

T 13
 A 3970
 ID ABG03970 standard; protein; 62 AA.

AC	ABG03970;
XX	
DT	13-FEB-2002 (first entry)
XX	
XX	Novel human diagnostic protein #3961.
DE	
DE	Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX	
KW	food supplement; medical imaging; diagnostic; genetic disorder.
KW	
XX	
OS	Homo sapiens.

XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PPF	30-MAR-2001; 2001WO-US008631.
XX	
PPR	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649167.
XX	
PPA	(HYSE-) HYSEQ INC.
XX	
PI	Dmanac RT, Liu C, Tang YT;
XX	
PFI	WPI; 2001-639362/73.
XX	
PFI	N-PSDB; AAS68157.
XX	
PPT	New isolated polynucleotide and encoded polypeptides, useful in
PPT	diagnostics, forensics, gene mapping, identification of mutations
PPT	responsible for genetic disorders or other traits and to assess
PPT	biodiversity.
PPT	

PS Claim 20; SEQ ID NO 34329; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG0377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 62 AA;

Query Match 40.2%; Score 82; DB 4; Length 62;
Best Local Similarity 42.9%; Pred. No. 0.073;
Matches 18; Conservative 9; Mismatches 11; Indels

QY	4 HLY----VSSKOKENISKENDVDLDEKEEAETEEEELEEK 41 : :: : : :: :: :
Db	7 HIYSNEVHXKBEETEEEEEETEEEEETEEEEEEEEEE 48

RESULT 14

ABG11244
ID ABG11244 standard; protein; 295 AA.

AC ABG11244;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #11235.

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.	XX
Homo sapiens.	OS

PN WO200175067-A2.

PD 11-OCT-2001.

30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

XX
XX
OUT REPORT / DATA

PI Drmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

DR N-PSDB; AAS75431.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

PS Claim 20: SEQ ID NO 41603; 103pp: English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and

Search completed: September 22, 2004, 18:06:24
Job time : 68.3967 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:03:03 ; Search time 19.314 Seconds
(without alignments)
109.592 Million cell updates/sec

Title: US-10-774-602-14
Perfect score: 204
Sequence: 1 MLSHLYVSKDKENISKEND.....VLDEKSEAEETEEELK 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	47.5	28	3	US-08-416-711-4
2	97	47.5	28	4	US-09-356-497-4
3	97	47.5	64	3	US-08-416-711-1
4	97	47.5	64	4	US-09-356-497-1
5	82	40.2	3135	1	US-08-323-1708-2
6	82	40.2	3135	4	US-08-954-441-2
7	79	38.7	714	2	US-08-990-114-3
8	79	38.7	714	4	US-09-241-333-3
9	77	37.7	740	1	US-08-257-073-5
10	76	37.3	905	2	US-08-574-959A-9
11	76	37.3	905	3	US-09-357-014-9
12	76	37.3	1135	2	US-08-574-959A-7
13	76	37.3	1135	3	US-09-357-014-7
14	75	36.8	739	3	US-09-022-983-2
15	73	35.8	111	4	US-09-886-319A-9
16	73	35.8	111	4	US-09-886-319A-10
17	73	35.8	1525	4	US-09-418-710-69
18	73	35.8	1912	4	US-09-495-714C-2
19	73	35.8	1977	4	US-09-495-714C-4
20	71.5	35.0	1312	4	US-09-345-882-29
21	70	34.3	66	4	US-09-621-976-5084
22	70	34.3	77	4	US-09-621-976-7521
23	70	34.3	1162	2	US-08-728-323A-2
24	70	34.3	1162	4	US-09-298-568-2
25	70	34.3	1182	4	US-09-410-399-2
26	69.5	34.1	357	1	US-08-552-142A-2
27	69.5	34.1	357	1	US-08-910-973-2

Sequence 2, Appli
Sequence 2, Appli
Sequence 27, Appli
Sequence 29, Appli
Sequence 24, Appli
Sequence 9, Appli
Sequence 8, Appli
Sequence 5087, Ap
Sequence 5083, Ap
Sequence 29, Appli
Sequence 4, Appli
Sequence 7, Appli
Sequence 214, App
Sequence 8, Appli
Sequence 14, Appli
Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-416-711-4
; Sequence 4, Application US/08416711
; Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: BOUHAROUN-TAYOUN, HASNAQ
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; TITLE OF INVENTION: PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-416-711-4

Query Match 47.5%; Score 97; DB 3; Length 28;
Best Local Similarity 95.0%; Pred. No. 0.00016;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
Db 9 MLSHLYVSSKDKENISKENE 28

RESULT 2
US-09-356-497-4
; Sequence 4, Application US/09356497
; Patent No. 6472519
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAO
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-356-497-4

Query Match 47.5%; Score 97; DB 4; Length 28;
Best Local Similarity 95.0%; Pred. No. 0.00016;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
Db 9 MLSHLYVSSKDKENISKENE 28

RESULT 3
US-08-416-711-1
; Sequence 1, Application US/08416711

Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAO
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-416-711-1

Query Match 47.5%; Score 97; DB 3; Length 64;
Best Local Similarity 95.0%; Pred. No. 0.00039;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
Db 45 MLSHLYVSSKDKENISKENE 64

RESULT 4
US-09-356-497-1
; Sequence 1, Application US/09356497
; Patent No. 6472519
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAO
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA

COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/356.497
FILING DATE: 19-JUL-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/416,711
FILING DATE: 08-AUG-1995
APPLICATION NUMBER: PCT/FR93/01024
FILING DATE: 18-OCT-1993
APPLICATION NUMBER: FR 92/12488
FILING DATE: 19-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-085-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 64 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-356-497-1

Query Match 47.5%; Score 97; DB 4; Length 64;
Best Local Similarity 95.0%; Pred. No. 0.00039;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLSHLYVSSKDKENISKEND 20
Db 45 MLSHLYVSSKDKENISKENE 64

RESULT 5
US-08-323-170B-2
Sequence 2, Application US/08323170B
Patent No. 5733772
GENERAL INFORMATION:
APPLICANT: Williamson, Kim C.
APPLICANT: Kaelow, David C.
TITLE OF INVENTION: Cloning and Expression of Plasmodium
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323,170B
FILING DATE: 13-OCT-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/010,409
FILING DATE: 29-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Einhorn, Gregory P.
REGISTRATION NUMBER: 38,440
REFERENCE/DOCKET NUMBER: 015280-113110US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0300
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-954-441-2

Query Match 40.2%; Score 82; DB 4; Length 3135;
Best Local Similarity 51.6%; Pred. No. 1.2;

ATTORNEY/AGENT INFORMATION:
NAME: Quine, Jonathan A.
REGISTRATION NUMBER: P-41,261
REFERENCE/DOCKET NUMBER: 015280-113100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-323-170B-2

Query Match 40.2%; Score 82; DB 1; Length 3135;
Best Local Similarity 51.6%; Pred. No. 1.2;
Matches 16; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Qy 11 DKENISKENDVLDKEKEEAETEEELK 41
Db 265 DEEDMSPRDNFVIDDEEEEEEEEEEE 295

RESULT 6
US-08-954-441-2
Sequence 2, Application US/08954441
Patent No. 6316000
GENERAL INFORMATION:
APPLICANT: Williamson, Kim C.
APPLICANT: Kaelow, David C.
TITLE OF INVENTION: Cloning and Expression of Plasmodium
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/954,441
FILING DATE: 20-OCT-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,170
FILING DATE: 13-OCT-1994
APPLICATION NUMBER: US 08/010,409
FILING DATE: 29-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Einhorn, Gregory P.
REGISTRATION NUMBER: 38,440
REFERENCE/DOCKET NUMBER: 015280-113110US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0300
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-954-441-2

Query Match 40.2%; Score 82; DB 4; Length 3135;
Best Local Similarity 51.6%; Pred. No. 1.2;


```
/
/ APPLICATION NUMBER: US/08/257,073
/ FILING DATE: 09-JUN-1994
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/075,783
/ FILING DATE: 11-JUN-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/852,305
/ FILING DATE: 18-MAR-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/672,183
/ FILING DATE: 20-MAR-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Frommer, William S.
/ REGISTRATION NUMBER: 25,506
/ REFERENCE/DOCKET NUMBER: 454310-2570
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 840-3333
/ TELEFAX: (212) 840-0712
/ TELEX: 425066 CURTMS
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 740 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FRAGMENT TYPE: internal
/ US-08-257-073-5

Query Match 37.7%; Score 77; DB 1; Length 740;
Best Local Similarity 50.0%; Pred. No. 0.86;
Matches 16; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 10 KDKENISKENDVLDKEKEAEETEEELK 41
Db 695 KEKEKEKEKEKEKEKEKEKEKEKE 726

RESULT 10
US-08-574-959A-9
; Sequence 9, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Joong, Ratna K. Vadlamudi
; and Jack L. Strominger
; TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,959A
; FILING DATE: 19-DEC-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 905 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-357-014-9

Query Match 37.3%; Score 76; DB 3; Length 905;
Best Local Similarity 40.0%; Pred. No. 1.4;
Matches 16; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

Qy 2 LSHLYVSSKDKENISKENDVLDKEKEAEETEEELK 41
Db 644 LTVININSDEEEEGEEEEEEEEEEEEEEEE 683

RESULT 11
US-09-357-014-9
; Sequence 9, Application US/09357014
; Patent No. 6291645
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Joong, Ratna K. Vadlamudi
; and Jack L. Strominger
; TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/357,014
; FILING DATE: 19-Jul-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/574,959
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 905 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-357-014-9

Query Match 37.3%; Score 76; DB 3; Length 905;
Best Local Similarity 40.0%; Pred. No. 1.4;
Matches 16; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

Qy 2 LSHLYVSSKDKENISKENDVLDKEKEAEETEEELK 41
Db 644 LTVININSDEEEEGEEEEEEEEEEEEEEEE 683

RESULT 12
US-08-574-959A-7
; Sequence 7, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
```


US-09-022-983-2

Query Match 36.8%; Score 75; DB 3; Length 739;
Best Local Similarity 50.0%; Pred. No. 1.4;
Matches 17; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
Qy 8 SSKDKENISKENDVDLDEKEEAEETEEELK 41
Db 436 TSKAETDDDDDDDDNEEEEEEEEEK 469

RESULT 15

US-09-886-319A-9
; Sequence 9, Application US/09886319A
; Patent No. 6586185
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Werner, Sabine
; APPLICANT: Halle, Jorn-Peter
; APPLICANT: Regenbogen, Johannes
; APPLICANT: Goppelt, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
; FILE REFERENCE: 50125/014002
; CURRENT APPLICATION NUMBER: US/09/886.319A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/222,081
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: DE 10030149.5
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-886-319A-9

Query Match 35.8%; Score 73; DB 4; Length 111;
Best Local Similarity 46.9%; Pred. No. 0.3;
Matches 15; Conservative 9; Mismatches 8; Indels 0; Gaps 0;
Qy 8 SSKDKENISKENDVDLDEKEEAEETEEELK 39
Db 38 NAQNEENGQADNEVDDEEEEGCGEEEEEE 69

Job completed: September 22, 2004, 18:11:03
Time : 20.314 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:03:03 ; Search time 13.1901 Seconds
(without alignments)
109.592 Million cell updates/sec

Title: US-10-774-602-13

Perfect score: 147

Sequence: 1 PEHKKEENMLSHLYVSSKDKENISKEND 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	97.3	28	3	US-08-416-711-4
2	143	97.3	28	4	US-09-356-497-4
3	143	97.3	64	3	US-08-416-711-1
4	143	97.3	64	4	US-09-356-497-1
5	48.5	33.0	652	1	US-08-261-663A-6
6	48.5	33.0	652	4	US-09-357-206A-5
7	48.5	33.0	652	4	US-09-813-742A-5
8	48.5	33.0	652	5	PCT-US95-07754A-6
9	48.5	33.0	746	4	US-09-134-001C-3214
10	47.5	32.3	68	4	US-09-621-976-7228
11	47.5	32.3	68	4	US-09-621-976-7229
12	47.5	32.3	68	4	US-09-621-976-7234
13	47.5	32.3	889	1	US-08-118-101A-4
14	47.5	32.3	1073	3	US-09-541-782-6
15	47.5	32.3	1073	4	US-09-723-820-6
16	47.5	32.3	1073	4	US-10-270-085-6
17	47.5	32.3	1177	4	US-09-134-001C-5106
18	47.5	32.3	10182	4	US-09-134-001C-3159
19	47	32.0	534	4	US-09-508-370A-5
20	46	31.3	27	3	US-08-416-711-3
21	46	31.3	27	4	US-09-356-497-3
22	46	31.3	181	4	US-09-134-001C-3897
23	46	31.3	516	4	US-09-496-320-13
24	45.5	31.0	239	4	US-09-513-783A-50
25	45.5	31.0	514	4	US-09-198-452A-214
26	45.5	31.0	541	4	US-09-513-783A-34
27	45.5	31.0	917	4	US-09-049-698-41

Sequence 6, Appli
Sequence 6428, Ap
Sequence 7850, Ap
Sequence 4697, Ap
Sequence 9, Appli
Sequence 4201, Ap
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 19, Appli
Sequence 4, Appli
Sequence 11, Appli
Sequence 8, Appli
Sequence 3818, Ap
Sequence 7803, Ap
Sequence 1045, Ap
Sequence 168, App

ALIGNMENTS

RESULT 1

US-08-416-711-4
; Sequence 4, Application US/08416711
; Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: BOUHAROUN-TAYOUN, HASNAQ
; APPLICANT: OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; TITLE OF INVENTION: PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-416-711-4

```

Query Match          97.3%; Score 143; DB 3; Length 28;
Best Local Similarity 96.4%; Pred. No. 9.2e-15;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKGENMLSHLYVSSKDKENISKEND 28
   |||||||||||||||||||||||||:
Db 1 PEHKGENMLSHLYVSSKDKENISKENE 28

RESULT 2
US-09-356-497-4
; Sequence 4, Application US/09356497
; Patent No. 6472519
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
;              BOUHAROUN-TAYOUN, HASNAQ
;              OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
;              PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;              P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
;
US-09-356-497-4

Query Match          97.3%; Score 143; DB 4; Length 28;
Best Local Similarity 96.4%; Pred. No. 9.2e-15;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKGENMLSHLYVSSKDKENISKEND 28
   |||||||||||||||||||||||||:
Db 1 PEHKGENMLSHLYVSSKDKENISKENE 28

RESULT 3
US-08-416-711-1
; Sequence 1, Application US/08416711

```

```

1 Patent No. 6017538
2 GENERAL INFORMATION:
3 APPLICANT: DRUILHE, PIERRE
4 APPLICANT: BOUHAROUN-TAYOUN, HASNAQ
5 APPLICANT: OEUVRAY, CLAUDE
6 TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
7 TITLE OF INVENTION: PROTECTIVE ANTIBODIES
8 NUMBER OF SEQUENCES: 10
9 CORRESPONDENCE ADDRESS:
10 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
11 ADDRESSEE: P.C.
12 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
13 CITY: ARLINGTON
14 STATE: VA
15 COUNTRY: USA
16 ZIP: 22202
17 COMPUTER READABLE FORM: disk
18 MEDIUM TYPE: Floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: Patent In Release #1.0, Version #1.30
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/08/416,711
24 FILING DATE: 08-AUG-1995
25 CLASSIFICATION: 424
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: PCT/FR93/01024
28 FILING DATE: 18-OCT-1993
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: FR 92/12488
31 FILING DATE: 19-OCT-1992
32 ATTORNEY/AGENT INFORMATION:
33 NAME: OBLON, NORMAN F.
34 REGISTRATION NUMBER: 24,618
35 REFERENCE/DOCKET NUMBER: 660-085-0 PCT
36 TELECOMMUNICATION INFORMATION:
37 TELEPHONE: 703-413-3000
38 TELEFAX: 703-413-2220
39 INFORMATION FOR SEQ ID NO: 1:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 64 amino acids
42 TYPE: amino acid
43 STRANDEDNESS: single
44 TOPOLOGY: linear
45 MOLECULE TYPE: peptide
46 US-08-416-711-1
47
48 Query Match 97.3%; Score 143; DB 3; Length 64;
49 Best Local Similarity 96.4%; Pred. NO. 2.5e-14;
50 Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
51
52 QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
53 |||||||||||||||||||||||||
54 Db 37 PEHKKEENMLSHLYVSSKDKENISKENE 64
55
56 RESULT 4
57 US-09-356-497-1
58 Sequence 1, Application US/09356497
59 Patent No. 6472519
60 GENERAL INFORMATION:
61 APPLICANT: DRUILHE, PIERRE
62 BOUHAROUN-TAYOUN, HASNAQ
63 OEUVRAY, CLAUDE
64 TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
65 PROTECTIVE ANTIBODIES
66 NUMBER OF SEQUENCES: 10
67 CORRESPONDENCE ADDRESS:
68 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
69 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
70 CITY: ARLINGTON
71 STATE: VA

```

```

1 Patent No. 6017538
2 GENERAL INFORMATION:
3 APPLICANT: DRUILHE, PIERRE
4 APPLICANT: BOUHAROUN-TAYOUN, HASNAQ
5 APPLICANT: OEUVRAY, CLAUDE
6 TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
7 TITLE OF INVENTION: PROTECTIVE ANTIBODIES
8 NUMBER OF SEQUENCES: 10
9 CORRESPONDENCE ADDRESS:
10 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
11 ADDRESSEE: P.C.
12 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
13 CITY: ARLINGTON
14 STATE: VA
15 COUNTRY: USA
16 ZIP: 22202
17 COMPUTER READABLE FORM: disk
18 MEDIUM TYPE: Floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: Patent In Release #1.0, Version #1.30
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/08/416,711
24 FILING DATE: 08-AUG-1995
25 CLASSIFICATION: 424
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: PCT/FR93/01024
28 FILING DATE: 18-OCT-1993
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: FR 92/12488
31 FILING DATE: 19-OCT-1992
32 ATTORNEY/AGENT INFORMATION:
33 NAME: OBLON, NORMAN F.
34 REGISTRATION NUMBER: 24,618
35 REFERENCE/DOCKET NUMBER: 660-085-0 PCT
36 TELECOMMUNICATION INFORMATION:
37 TELEPHONE: 703-413-3000
38 TELEFAX: 703-413-2220
39 INFORMATION FOR SEQ ID NO: 1:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 64 amino acids
42 TYPE: amino acid
43 STRANDEDNESS: single
44 TOPOLOGY: linear
45 MOLECULE TYPE: peptide
46 US-08-416-711-1
47
48 Query Match 97.3%; Score 143; DB 3; Length 64;
49 Best Local Similarity 96.4%; Pred. NO. 2.5e-14;
50 Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
51
52 QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
53 |||||||||||||||||||||||||
54 Db 37 PEHKKEENMLSHLYVSSKDKENISKENE 64
55
56 RESULT 4
57 US-09-356-497-1
58 Sequence 1, Application US/09356497
59 Patent No. 6472519
60 GENERAL INFORMATION:
61 APPLICANT: DRUILHE, PIERRE
62 BOUHAROUN-TAYOUN, HASNAQ
63 OEUVRAY, CLAUDE
64 TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
65 PROTECTIVE ANTIBODIES
66 NUMBER OF SEQUENCES: 10
67 CORRESPONDENCE ADDRESS:
68 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
69 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
70 CITY: ARLINGTON
71 STATE: VA

```

COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/356.497
 FILING DATE: 19-Jul-1999
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/416,711
 FILING DATE: 08-AUG-1995
 APPLICATION NUMBER: PCT/FR93/01024
 FILING DATE: 18-OCT-1993
 APPLICATION NUMBER: FR 92/12488
 FILING DATE: 19-OCT-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 660-085-0 PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 64 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-356-497-1
 Query Match 97.3%; Score 143; DB 4; Length 64;
 Best Local Similarity 96.4%; Pred. No. 2.5e-14;
 Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PEHKEENMLSHLYVSSKDKENISKEND 28
 DB 37 PEHKEENMLSHLYVSSKDKENISKENE 64
 RESULT 5
 US-08-261-663A-6
 Sequence 6, Application US/08261663A
 Patent No. 5571706
 GENERAL INFORMATION:
 APPLICANT: Baker, Barbara J
 APPLICANT: Whitham, Steven A
 TITLE OF INVENTION: Plant Virus Resistance Gene and Methods
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Margaret A. Connor, USDA-ARS
 STREET: 800 Buchanan Street
 CITY: Albany
 STATE: CA
 COUNTRY: USA
 ZIP: 94710
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/261.663A
 FILING DATE:
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Connor, Margaret A
 REGISTRATION NUMBER: 30043
 REFERENCE/DOCKET NUMBER: 0094.94

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (510) 559-6067
 TELEFAX: (510) 559-5777
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 652 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-261-663A-6
 Query Match 33.0%; Score 48.5; DB 1; Length 652;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 11; Conservative 3; Mismatches 7; Indels 1; Gaps 1;
 QY 8 NMLSHLYVSSKDKENIS-KEND 28
 DB 606 NSLRHLWTETKKNIAKEGD 627
 RESULT 6
 US-09-357-206A-5
 Sequence 5, Application US/09357206A
 Patent No. 6372962
 GENERAL INFORMATION:
 APPLICANT: Dinesh-Kumar, S.
 APPLICANT: Baker, Barbara
 TITLE OF INVENTION: Pathogen Resistance in Plants using cdNA-N/Intron Constructs
 FILE REFERENCE: 042250/191805 (5830-5)
 CURRENT APPLICATION NUMBER: US/09/357,206A
 CURRENT FILING DATE: 1999-07-20
 PRIOR APPLICATION NUMBER: US 60/093,494
 PRIOR FILING DATE: 1998-07-20
 NUMBER OF SEQ ID NOS: 22
 SOFTWARE: Patentin version 3.0
 SEQ ID NO 5
 LENGTH: 652
 TYPE: PRT
 ORGANISM: Nicotiana glutinosa
 US-09-357-206A-5
 Query Match 33.0%; Score 48.5; DB 4; Length 652;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 11; Conservative 3; Mismatches 7; Indels 1; Gaps 1;
 QY 8 NMLSHLYVSSKDKENIS-KEND 28
 DB 606 NSLRHLWTETKKNIAKEGD 627
 RESULT 7
 US-09-813-742A-5
 Sequence 5, Application US/09813742A
 Patent No. 6630618
 GENERAL INFORMATION:
 APPLICANT: Baker, Barbara J
 APPLICANT: Dinesh-Kumar, S. P.
 TITLE OF INVENTION: NON-PATHOGEN INDUCED SYSTEMIC ACQUIRED RESISTANCE (SAR) IN PLANTS
 FILE REFERENCE: 42250/209601 (5830-12)
 CURRENT APPLICATION NUMBER: US/09/813,742A
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,027
 PRIOR FILING DATE: 2000-03-21
 NUMBER OF SEQ ID NOS: 11
 SOFTWARE: Patentin version 3.1
 SEQ ID NO 5
 LENGTH: 652
 TYPE: PRT
 ORGANISM: Nicotiana glutinosa
 US-09-813-742A-5
 Query Match 33.0%; Score 48.5; DB 4; Length 652;
 Best Local Similarity 50.0%; Pred. No. 48;

Matches	11;	Conservative	3;	Mismatches	7;	Indels	1;	Gaps	1;
---------	-----	--------------	----	------------	----	--------	----	------	----

Qy	8	NMLSHLYVSSKDKENIS-KEND	28
		:	
Db	606	NSLRHLWTETKKKNNAIEKGD	627

RESULT 8

```

PCT-US95-07754A-6
; Sequence 6, Application PC/TUS9507754A
; GENERAL INFORMATION:
; APPLICANT: Baker, Barbara J
; APPLICANT: Whitham, Steven A
; TITLE OF INVENTION: Plant Virus Resistance Gene and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Margaret A. Connor, USDA-ARS
; STREET: 800 Buchanan Street
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07754A

```

```

/ CLASSIFICATION:
/
/ ATTORNEY/AGENT INFORMATION:
/
/   NAME: Connor, Margaret A
/
/   REGISTRATION NUMBER: 30043
/
/   REFERENCE/DOCKET NUMBER: 0094, 94
/
/ TELECOMMUNICATION INFORMATION:
/
/   TELEPHONE: (510) 559-6067
/
/   TELEFAX: (510) 559-5777
/
/   INFORMATION FOR SEQ ID NO: 6:
/     SEQUENCE CHARACTERISTICS:
/       LENGTH: 652 amino acids
/       TYPE: amino acid
/       TOPOLOGY: linear
/
/     MOLECULE TYPE: protein
/
/ PCT-US95-07754A-6

```

Query Match 33.0%; Score 48.5; DB 5; Length 652;
 Percent Local Similarity 50.0%; Pred. No. 48;
 Matches 11; Conservative 3; Mismatches 7; Indels

Qy	8	NMLSHLYVSSKDKENIS-KEND	28
		: :	
Db	606	NSLRHLWTETKKONNIAEKEGD	627

RESULT 9

```

US-09-134-001C-3214
; Sequence 3214, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: LYNN Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GPC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3214
; LENGTH: 746

```

; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3214

Query Match 33.0%; Score 48.5; DB 4; Length 746;
Best Local Similarity 42.9%; Pred. No. 56;
Matches 12; Conservative 5; Mismatches 8; Indels

Qy 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
|::|||:|||||:

Dp 645 POYLIEDWMS---ISSKGEDFKKPND 669

RESULT 10

```

US-09-621-976-7228
; Sequence 7228, Application US/09621976
; Patent No. 6639083
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: EGTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7228
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7228

```

Query Match 32.3%; Score 47.5; DB 4; Length 68;
Best Local Similarity 45.8%; Pred. No. 4.5;
Matches 11; Conservative 5; Mismatches 3; Indels

QY 5 KEENMLSHLYVSSKDKENISKEND 28
||| |::| | ||::|||
Db 4 KEERALNNLIV-----ENNVOEND 22

RESULT 11

```

US-09-621-976-7229
; Sequence 7229, Application US/09621976
; Patent No. 6639083
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7229
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7229

```

Query Match 32.3%; Score 47.5; DB 4; Length 68;
Best Local Similarity 45.8%; Pred. No. 4.5;
Matches 11: Conservative 5; Mismatches 3; Indels

QY 5 KEENMLSHLYVSSKDKENISKEND 28
||| | : | | ||| : |||
Db 4 KEERALNNLIV-----ENNVOEND 22

RESULT 12
US-09-621-976-7234

```

Qy 6 ENNLSHLVYSSKDKENISKEND 28
   ||| ||| : ||| : |||
Db 685 EQSGL-LYDSGQDENITHEDD 706

RESULT 14
US-09-541-782-6
; Sequence 6, Application US/09541782
; Patent No. 6284480
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/541,782
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: S.pombe
US-09-541-782-6

Query Match 32.3%; Score 47.5; DB 3; Length 1073;
Best Local Similarity 42.9%; Pred. No. 1.2e+02;
Matches 12; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

Qy 2 EHKKE-ENNLSHLVSSKDKENISKEND 28
   ||||| : ||| : ||| : |||
Db 499 EHKKEVALQLVNSSTELSVKSENE 526

RESULT 15
US-09-723-820-6
; Sequence 6, Application US/09723820
; Patent No. 6468760
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/723,820
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/541,782
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: S.pombe
US-09-723-820-6

Query Match 32.3%; Score 47.5; DB 4; Length 1073;
Best Local Similarity 42.9%; Pred. No. 1.2e+02;
Matches 12; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

Qy 2 EHKKE-ENNLSHLVSSKDKENISKEND 28
   ||||| : ||| : ||| : |||
Db 499 EHKKEVALQLVNSSTELSVKSENE 526

Search completed: September 22, 2004, 18:11:02
Job time : 14.1901 secs

```



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:09:13 ; Search time 60.3967 Seconds
(without alignments)
148.878 Million cell updates/sec

Title: US-10-774-602-13

Perfect score: 147

Sequence: 1 PEHKKEENMLSHLYVSSKOKENISKEND 28

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Maximum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147	100.0	28	14	US-10-294-770-13
2	147	100.0	28	16	US-10-774-602-13
3	143	97.3	28	14	US-10-294-770-4
4	143	97.3	28	16	US-10-238-741-4
5	143	97.3	28	16	US-10-774-602-4
6	143	97.3	64	14	US-10-294-770-1
7	143	97.3	64	16	US-10-238-741-1
8	143	97.3	64	16	US-10-774-602-1
9	101	68.7	41	14	US-10-294-770-14
10	101	68.7	41	16	US-10-774-602-14
11	54	36.1	199	16	US-10-437-963-161536
12	53	36.1	92	16	US-10-437-963-109657
13	53	36.1	426	9	US-09-731-872-310
14	53	36.1	426	9	US-09-731-872-317
15	53	36.1	426	10	US-09-876-997-310

16	53	36.1	426	10	US-09-876-997-317	Sequence 317, App
17	53	36.1	426	16	US-10-655-601-5	Sequence 5, Appli
18	53	36.1	481	9	US-09-731-872-415	Sequence 415, App
19	53	36.1	481	10	US-09-876-997-415	Sequence 415, App
20	52.5	35.7	307	16	US-10-437-963-140346	Sequence 140346, A
21	52.5	35.7	654	12	US-10-425-114-65105	Sequence 65105, A
22	52	35.4	81	12	US-10-424-599-150366	Sequence 150366, A
23	52	35.4	443	12	US-10-282-122A-48757	Sequence 48757, A
24	52	35.4	749	15	US-10-369-493-21933	Sequence 21933, A
25	51.5	35.0	1182	12	US-10-282-122A-53445	Sequence 53445, A
26	51	34.7	98	16	US-10-437-963-159924	Sequence 159924, A
27	51	34.7	174	14	US-10-238-075-558	Sequence 558, App
28	51	34.7	204	16	US-10-767-701-31636	Sequence 31636, A
29	51	34.7	415	12	US-10-282-122A-70442	Sequence 70442, A
30	51	34.7	572	12	US-10-267-503-235	Sequence 235, App
31	51	34.7	572	16	US-10-287-226-661	Sequence 661, App
32	51	34.7	1441	16	US-10-437-963-159919	Sequence 159919, A
33	50.5	34.4	258	12	US-10-424-599-246340	Sequence 246340, A
34	50	34.0	515	12	US-10-282-122A-53329	Sequence 53329, A
35	49	33.3	209	16	US-10-437-963-117713	Sequence 117713, A
36	49	33.3	231	12	US-10-425-114-68769	Sequence 68769, A
37	49	33.3	344	12	US-10-425-114-45399	Sequence 45399, A
38	49	33.3	405	12	US-10-425-114-68612	Sequence 68612, A
39	49	33.3	473	15	US-10-094-749-1936	Sequence 1936, App
40	49	33.3	473	16	US-10-408-765A-2491	Sequence 2491, App
41	49	33.3	680	16	US-10-437-963-123028	Sequence 123028, A
42	49	33.3	741	16	US-10-437-963-117706	Sequence 117706, A
43	48.5	33.0	652	9	US-09-813-742-5	Sequence 5, Appli
44	48.5	33.0	743	12	US-10-282-122A-70586	Sequence 70586, A
45	48.5	33.0	1700	16	US-10-437-963-162893	Sequence 162893, A

ALIGNMENTS

RESULT 1
US-10-294-770-13
; Sequence 13, Application US/10294770
; Publication No. US20030161840A1
; GENERAL INFORMATION:
; APPLICANT: DRUIHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759USOCIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-294-770-13

Query Match 100.0%; Score 147; DB 14; Length 28;

Best Local Similarity 100.0%; Pred. No. 1,1e-12; Indels 0; Gaps 0;

Matches 28; Conservative 0; Mismatches 0;

QY 1 PEHKKEENMLSHLYVSSKOKENISKEND 28

Db 1 PEHKKEENMLSHLYVSSKOKENISKEND 28

RESULT 2

US-10-774-602-13

; Sequence 13, Application US/10774602

```

; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791USODIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; TYPE: PRT
; LENGTH: 28
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-774-602-13

```

```

Query Match      100.0%; Score 147; DB 16; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PEHKEENMLSHLYVSSKDKENISKEND 28
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 PEHKEENMLSHLYVSSKDKENISKEND 28

```

```

RESULT 3
US-10-294-770-4
; Sequence 4, Application US/10294770
; Publication No. US20030161840A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759USOCIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-294-770-4

```

```

Query Match      97.3%; Score 143; DB 14; Length 28;
Best Local Similarity 96.4%; Pred. No. 3.9e-12;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PEHKEENMLSHLYVSSKDKENISKEND 28
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 PEHKEENMLSHLYVSSKDKENISKENE 28

```

```

RESULT 4
US-10-238-741-4
; Sequence 4, Application US/10238741
; Publication No. US20040096466A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE

```

```

; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/238,741
; FILING DATE: 09-Nov-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-238-741-4

```

```

Query Match      97.3%; Score 143; DB 16; Length 28;
Best Local Similarity 96.4%; Pred. No. 3.9e-12;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PEHKEENMLSHLYVSSKDKENISKEND 28
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 PEHKEENMLSHLYVSSKDKENISKENE 28

```

```

RESULT 5
US-10-774-602-4
; Sequence 4, Application US/10774602
; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791USODIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08

```

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28

Db 37 PEHKEENMLSHLYVSSKDKENISKENE 64
|||||

RESULT 9
US-10-294-770-14
; Sequence 14, Application US/10294770
; Publication No. US20030161840A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759US0CIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 14
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-294-770-14

Query Match 68.7%; Score 101; DB 14; Length 41;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 MLSHLYVSSKDKENISKEND 28
|||||

Db 1 MLSHLYVSSKDKENISKEND 20
|||||

RESULT 10
US-10-774-602-14
; Sequence 14, Application US/10774602
; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791US0DIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 14
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-774-602-14

Query Match 68.7%; Score 101; DB 16; Length 41;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 MLSHLYVSSKDKENISKEND 28
|||||

Db 1 MLSHLYVSSKDKENISKEND 20

RESULT 11
US-10-437-963-161536
; Sequence 161536, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 161536
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_60710C.1.pep
US-10-437-963-161536

Query Match 36.7%; Score 54; DB 16; Length 199;
Best Local Similarity 33.3%; Pred. No. 27;
Matches 9; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 2 EHKKEENMLSHLYVSSKDKENISKEND 28
|||||

Db 131 EHEENQKRLQLKLNKDIESLKQND 157
|||||

RESULT 12
US-10-437-963-109657
; Sequence 109657, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 109657
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13794C.1.pep
US-10-437-963-109657

Query Match 36.1%; Score 53; DB 16; Length 92;
Best Local Similarity 39.1%; Pred. No. 16;
Matches 9; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 4 KKEENMLSHLYVSSKDKENISKE 26
|||||

Db 57 KFKDTMLEHLVLEAKDKKTINRQ 79
|||||

RESULT 13
US-09-731-872-310
; Sequence 310, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US3 REG
; CURRENT APPLICATION NUMBER: US/09/731.872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 310
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -28..-1
US-09-731-872-310

Query Match 36.1%; Score 53; DB 9; Length 426;
Best Local Similarity 33.3%; Pred. No. 86;
Matches 13; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

Qy 2 EHKKEENMLSH-----LYVSSKDKENISKE 26
Db 134 QHRKEEKMKSHKLEEFWLNKKSEVLYYTVKKGNISQ 172

RESULT 14
US-09-731-872-317
; Sequence 317, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US3 REG
; CURRENT APPLICATION NUMBER: US/09/731.872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 317
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -28..-1
US-09-731-872-317

Query Match 36.1%; Score 53; DB 9; Length 426;
Best Local Similarity 33.3%; Pred. No. 86;
Matches 13; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

Qy 2 EHKKEENMLSH-----LYVSSKDKENISKE 26
Db 134 QHRKEEKMKSHKLEEFWLNKKSEVLYYTVKKGNISQ 172

RESULT 15
US-09-876-997-310
; Sequence 310, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US4 CIP
; CURRENT APPLICATION NUMBER: US/09/876.997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731.872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 310
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -28..-1
US-09-876-997-310

Query Match 36.1%; Score 53; DB 10; Length 426;
Best Local Similarity 33.3%; Pred. No. 86;
Matches 13; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

Qy 2 EHKKEENMLSH-----LYVSSKDKENISKE 26
Db 134 QHRKEEKMKSHKLEEFWLNKKSEVLYYTVKKGNISQ 172

Search completed: September 22, 2004, 18:34:05
Job time : 61.3967 secs

Miss Page Blank (1891)

Result No.	Query	Score	Match	Length	DB	ID	Description
1	59	47.6	890	2	S22452	surface exclusion	
2	55.5	44.8	820	2	T51510	hypothetical prote	
3	55	44.4	752	2	H6770	hypothetical prote	
4	55	44.4	891	2	G1662	130K surface exclu	
5	54.5	44.0	8162	2	T51040	hypothetical prote	
6	53	42.7	463	2	T13425	regulatory protein	
7	52	41.9	213	2	T22984	hypothetical prote	
8	52	41.9	420	2	T48234	hypothetical prote	
9	50	40.3	367	2	B82363	glpE protein VC010	
10	49.5	39.9	347	2	B85432	hypothetical prote	
11	49	39.5	180	2	F97149	probable flavodoxi	
12	49	39.5	651	2	T15624	hypothetical prote	
13	48	38.7	988	2	A24341	transposase - Pseu	
14	47.5	38.3	100	1	R3EC31	ribosomal protein	
15	47.5	38.3	100	2	G91151	50S ribosomal subu	
16	47.5	38.3	100	2	C85997	50S ribosomal subu	
17	47.5	38.3	174	2	G70220	exported protein A	
18	47.5	38.3	1131	2	A49393	activator 1 large	
19	47	37.9	51	2	F97708	hypothetical prote	
20	47	37.9	67	2	T14410	pollen coat protei	
21	47	37.9	119	2	E81939	probable 50S ribos	
22	47	37.9	119	2	F81165	50S ribosomal prot	
23	47	37.9	179	2	T42293	hypothetical prote	
24	47	37.9	441	2	B96953	methyl-accepting c	
25	47	37.9	1240	2	T04833	hypothetical prote	
26	47	37.9	1386	2	AC1533	surface protein (L	
27	46.5	37.5	100	2	AC1006	50S ribosomal chai	
28	46.5	37.5	482	2	AF1120	conserved hypothet	
29	46	37.1	338	2	E69964	NADH-dependent fla	

```
Db      446 YEKSNAYQ---EAIKAEAS 464
      ||| ||| ||| :|: ||||
      446 YEKSNAYQ---EAIKAEAS 464

RESULT 3
H86770
hypothetical protein yliB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: H86770
R:Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: H86770
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-752 <STO>
A:Cross-references: GB:AE005176; PID:g12724133; PIDN:AAK05266.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yliB

Query Match      44.4%; Score 55; DB 2; Length 752;
Best Local Similarity 52.6%; Pred. No. 9.1;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      2 EKAKNAYQKANOAVLKAKE 20
      ||||| ||| :|: ||||
      450 EEAKNTYHFANEMILKSKK 468

Db      450 EEAKNTYHFANEMILKSKK 468

RESULT 4
G41662
130K surface exclusion protein Sec10 precursor - Enterococcus faecalis plasmid pCF10
C:Species: Enterococcus faecalis
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
C:Accession: G41662
R:Xao, S.M.; Olmsted, S.B.; Viksins, A.S.; Gallo, J.C.; Dunny, G.M.
J. Bacteriol. 173, 7650-7664, 1991
A:Title: Molecular and genetic analysis of a region of plasmid pCF10 containing positive
terococcus faecalis.
A:Reference number: A41662; MUID:92041679; PMID:1938961
A:Accession: G41662
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-891 <KAO>
A:Cross-references: GB:M64978; NID:g150552; PIDN:AAA65847.1; PID:g150554
C:Genetics:
A:Genome: plasmid

Query Match      44.4%; Score 55; DB 2; Length 891;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      1 YEKAKNAYQKANOAVLKAKEA 21
      ||||| ||| ||| :|: ||||
      709 YEKALNELNKAAEAAVVOAQA 729

Db      709 YEKALNELNKAAEAAVVOAQA 729

RESULT 5
T51040
hypothetical protein B15120.100 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T51040
R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51040
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1162 <SCH>
```

```
A:Cross-references: EMBL:AL389900; GSPDB:GN00116; NCSP:B15120.100
A:Experimental source: BAC clone B15120; strain OR74A
C:Genetics:
A:Gene: NCSP:B15120.100
A:Map position: 6
A:Introns: 59/2; 135/3; 171/3; 309/1; 430/2

Query Match      44.0%; Score 54.5; DB 2; Length 1162;
Best Local Similarity 59.3%; Pred. No. 16;
Matches 16; Conservative 3; Mismatches 3; Indels 5; Gaps 2;

QY      2 EKA-KNAYQ----KANQAVLKAKEASS 23
      ||||| :|: ||| ||||| :|:
      632 EKALNKAQEMARIKAEQAALKAKEANA 658

Db      632 EKALNKAQEMARIKAEQAALKAKEANA 658

RESULT 6
T13425
regulatory protein K10, oocyte-specific - fruit fly (Drosophila melanogaster)
N:Alternate names: protein EG:30B8.5
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 15-Oct-1999
C:Accession: T13425; A28826
R:Murphy, L.; Harris, D.; Barrell, B.
submitted to the EMBL Data Library, April 1999
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A:Reference number: Z17668
A:Accession: T13425
A>Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-463 <MUR>
A:Cross-references: EMBL:AL009195; NID:el355203; PID:el202207; PIDN:CAA15702.1
Genes Dev. 2, 891-900, 1988
R:Prost, E.; Deryckere, F.; Roos, C.; Haenlin, M.; Pantescio, V.; Mohler, E.
A:Title: Role of the oocyte nucleus in determination of the dorsoventral polarity of Dros
A:Reference number: A28826
A:Accession: A28826
A:Molecule type: mRNA
A:Residues: 1-216, 'HH', 279-281, 'VDHHR', 287-463 <PRO>
A:Cross-references: GB:X12836; NID:g8148; PID:g295771
C:Genetics:
A:Gene: FlyBase:fs(1)K10
A:Cross-references: FlyBase:FBgn0000810
A:Map position: X
A:Introns: 432/3
C:Keywords: DNA binding

Query Match      42.7%; Score 53; DB 2; Length 463;
Best Local Similarity 47.8%; Pred. No. 11;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY      1 YEKAKNAYQKANOAVLKAKEASS 23
      ||||| :|: ||| ||||| :|:
      414 YEAAREYLKQEAATVKAKDAKS 436

Db      414 YEAAREYLKQEAATVKAKDAKS 436

RESULT 7
T22984
hypothetical protein F59B8.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22984
R:Matthews, P.
submitted to the EMBL Data Library, January 1996
A:Reference number: Z19647
A:Accession: T22984
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-213 <WIL>
A:Cross-references: EMBL:268343; PIDN:CAA92779.1; GSPDB:GN00022; CESP:F59B8.1
A:Experimental source: clone F59B8
C:Genetics:
A:Gene: CESP:F59B8.1
```


A;Map position: 4
A;Introns: 31/2; 65/3; 94/2; 138/1; 170/3

Query Match 41.9%; Score 52; DB 2; Length 213;
Best Local Similarity 43.5%; Pred. No. 6.8;
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 1 YEKAKNAYQKANOAVLKAKKASS 23
| : ||| : : ||| : : ||| : :
Db 30 YRLKNAKXKQEDYLSKEANA 52

RESULT 8

T48234
hypothetical protein T7H20.180 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 05-May-2000
C;Accession: T48234
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24488
A;Accession: T48234
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-420 <BEV>
A;Cross-references: EMBL:AL162508
A;Experimental source: cultivar Columbia; BAC clone T7H20
C;Genetics:
A;Map position: 5
A;Introns: 67/1; 109/2; 134/1; 149/3; 196/1; 211/1; 274/2; 295/1; 328/3; 347/1; 362/1
A;Note: T7H20.180
C;Superfamily: Arabidopsis thaliana hypothetical protein T7H20.180

Query Match 41.9%; Score 52; DB 2; Length 420;
Best Local Similarity 52.4%; Pred. No. 13;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 YEKAKNAYQKANOAVLKAKKA 21
| : ||| : : ||| : : ||| : :
Db 227 FELAKEMYQKAIQGVTTKES 247

RESULT 9

B82363
glpE protein VC0100 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: B82363
R;Maidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Rosen, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.
R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406: 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: B82363
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-106 <HEI>
A;Cross-references: GB:AE004101; GB:AE003852; NID:g9654497; PIDN:AAF93278.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0100
A;Map position: 1
C;Superfamily: glpE protein

Query Match 40.3%; Score 50; DB 2; Length 106;
Best Local Similarity 40.9%; Pred. No. 6.4;
Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 4 AKNAYQKANOAVLKAKKASSYD 25
| : ||| : : ||| : : ||| : :
Db 37 AKNAYHLTNSQVQFMEQAEFD 58

RESULT 10

B85432
hypothetical protein AT4g36600 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: B85432
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402: 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: B85432
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-347 <STO>
A;Cross-references: GB:NC_001368; NID:g7270608; PIDN:CAB80326.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g36600
A;Map position: 4

Query Match 39.9%; Score 49.5; DB 2; Length 347;
Best Local Similarity 36.8%; Pred. No. 24;
Matches 14; Conservative 4; Mismatches 7; Indels 13; Gaps 1;

Qy 1 YEKAKNAYQKANOAVLKKA-----KEASSYD 25
| : ||| : : ||| : : ||| : :
Db 145 YDKAGSAYEKAGQAKDMAYDKAGQAKDMAYDKVGSAYD 182

RESULT 11

F97149
probable flavodoxin [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: F97149
R;Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J. Bacteriol. 183: 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: F97149
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-180 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK79985.1; PID:g15025010; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2026

Query Match 39.5%; Score 49; DB 2; Length 180;
Best Local Similarity 58.8%; Pred. No. 15;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 6 NAYQKANOAVLKAKKAS 22
| : ||| : : ||| : : ||| : :
Db 28 NKPRQANQKILKAKKAN 44

RESULT 12

T15624
hypothetical protein C25H3.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15624
R;Johnson, D.
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of C. elegans cosmid C25H3.
A;Reference number: Z18379
A;Accession: T15624
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-651 <JOH>
A;Cross-references: EMBL:U29535; NID:g868251; PID:g868256; PIDN:AAA68785.1; CESP:C25H3.6

Query Match 40.3%; Score 50; DB 2; Length 106;
Best Local Similarity 40.9%; Pred. No. 6.4;
Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 4 AKNAYQKANOAVLKAKKASSYD 25
| : ||| : : ||| : : ||| : :
Db 37 AKNAYHLTNSQVQFMEQAEFD 58

A;Experimental source: strain Bristol N2
 C;Genetics:
 A;Gene: CESP:C25H3.6
 A;Introns: 58/1; 213/3; 273/1; 446/2; 468/2; 484/2; 510/1; 552/1; 609/1

Query Match 39.5%; Score 49; DB 2; Length 651;
 Best Local Similarity 40.0%; Pred. No. 53;
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 6 NAYKANQAVLKKEASSYD 25
 | : | | : : : | |
 Db 526 NSQSFNMAIVAPQAPDYD 545

RESULT 13
 A24341
 transposase - Pseudomonas aeruginosa transposon Tn501
 C;Species: Pseudomonas aeruginosa
 C;Date: 29-Aug-1987 #sequence_revision 29-Aug-1987 #text_change 21-Jul-2000
 C;Accession: A24341
 R;Brown, N.L.; Winnie, J.N.; Fritzinger, D.; Pridmore, R.D.
 Nucleic Acids Res. 13, 5657-5669, 1985
 A;Title: The nucleotide sequence of the trpA gene completes the sequence of the Pseudomonas
 A;Reference number: A24341; MUID:85297764; PMID:2994007
 A;Accession: A24341
 A;Molecule type: DNA
 A;Residues: 1-988 <BRO>
 A;Cross-references: GB:Z00027; GB:K00031; GB:K01725; GB:X01297; GB:X03406; NID:g43714; F
 C;Genetics:
 A;Gene: tnpA
 C;Superfamily: transposase Tn21
 C;Keywords: DNA binding

Query Match 38.7%; Score 48; DB 2; Length 988;
 Best Local Similarity 62.5%; Pred. No. 1.le+02;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 YOKANQAVLKKEASS 23
 | : | | : : : | |
 Db 330 YSKIGQALLLEKEAGS 345

RESULT 14
 R5EC23
 ribosomal protein L23 [validated] - Escherichia coli (strain K-12)
 C;Species: Escherichia coli
 C;Date: 28-Feb-1980 #sequence_revision 31-Oct-1997 #text_change 01-Mar-2002
 C;Accession: A65125; D23129; A02814
 R;Wattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503
 A;Accession: A65125
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-100 <BLAT>
 A;Cross-references: GB:AE000408; GB:U00096; NID:g1789694; PIDN:AACT6343.1; PID:g1789714;
 A;Experimental source: strain K-12, substrain MGL655
 R;Zurawski, G.; Zurawski, S.M.
 Nucleic Acids Res. 13, 4521-4526, 1985
 A;Title: Structure of the Escherichia coli S10 ribosomal protein operon.
 A;Reference number: A23129; MUID:85242118; PMID:3892488
 A;Accession: D23129
 A;Molecule type: DNA
 A;Residues: 1-100 <ZUR>
 A;Cross-references: GB:X02613; NID:g42825; PIDN:CAA26462.1; PID:g534975
 A;Note: translation of initiator Met is not shown; the initiator overlaps the last codon
 R;Wittmann-Liebold, B.; Greuer, B.
 FEBS Lett. 108, 69-74, 1979
 A;Title: Primary structure of protein L23 from the Escherichia coli ribosome.
 A;Reference number: A02814; MUID:80092111; PMID:391594
 A;Accession: A02814

A;Molecule type: protein
 A;Residues: 1-79, 81-100 <WIT>
 A;Experimental source: strain K12
 R;Arnold, R.J.; Reilly, J.P.
 Anal. Biochem. 269, 105-112, 1999
 A;Title: Observation of Escherichia coli ribosomal proteins and their posttranslational n
 A;Reference number: A59071; MUID:99196679; PMID:10094780
 A;Content: annotation; mass spectrographic analysis
 A;Note: no post-translational modifications were observed in mass spectrographic analysis
 C;Genetics:
 A;Gene: rplW
 A;Map position: 73 min
 C;Complex: the ribosome is composed of the large (50S) and small (30S) subunit; the large
 S rRNA and 22 distinct proteins
 C;Complex: large subunit ribosomal proteins: L1 (PIR:R5EC1), L3 (PIR:R5EC3), L2 (PIR:R5EC
 (PIR:R5EC11), L13 (PIR:R5EC13), L14 (PIR:R5EC14), L15 (PIR:R5EC15), L16 (PIR:R5EC16), L1;
 C;Function:
 A;Pathway: protein biosynthesis
 C;Superfamily: Escherichia coli ribosomal protein L23
 C;Keywords: protein biosynthesis; ribosome
 F;1-100/Product: ribosomal protein L23 #status experimental <MAT>

Query Match 38.3%; Score 47.5; DB 1; Length 100;
 Best Local Similarity 54.5%; Pred. No. 13;
 Matches 12; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 2 EKAKNAYKANQAVLK-AKEAS 22
 | | | : | : | | | | : | : |
 Db 18 EKASTAMEKSNITVLKVKANDAT 39

RESULT 15
 G91151
 50S ribosomal subunit protein L23 [imported] - Escherichia coli (strain O157:H7, substra
 C;Species: Escherichia coli
 C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C;Accession: G91151
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
 A;Reference number: A99629; MUID:21156231; PMID:11258796
 A;Accession: G91151
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-100 <HAY>
 A;Cross-references: GB:BA000007; PIDN:BA837606.1; PID:g13363656; GSPDB:GN00154
 A;Experimental source: strain O157:H7, substrain RIMD 0509952
 C;Genetics:
 A;Gene: ECs4183
 C;Superfamily: Escherichia coli ribosomal protein L23

Query Match 38.3%; Score 47.5; DB 2; Length 100;
 Best Local Similarity 54.5%; Pred. No. 13;
 Matches 12; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 2 EKAKNAYKANQAVLK-AKEAS 22
 | | | : | : | | | | : | : |
 Db 18 EKASTAMEKSNITVLKVKANDAT 39

Search completed: September 22, 2004, 18:09:58
 Job time : 13.5041 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:03:03 ; Search time 11.7769 Seconds
(without alignments)
109.592 Million cell updates/sec

Title: US-10-774-602-11

Perfect score: 124

Sequence: 1 YERAKNAYKQNAQVLKAKEASSYD 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	93.5	64	3	US-08-416-711-1
2	116	93.5	64	4	US-09-356-497-1
3	96	77.4	23	3	US-08-416-711-2
4	96	77.4	23	4	US-09-356-497-2
5	50	40.3	28	1	US-08-303-025-12
6	50	40.3	28	2	US-08-436-703B-1
7	50	40.3	29	1	US-08-152-488-10
8	50	40.3	29	1	US-08-152-488-11
9	50	40.3	29	1	US-08-303-025-10
10	50	40.3	29	1	US-08-303-025-11
11	50	40.3	29	1	US-08-303-025-13
12	50	40.3	29	1	US-08-677-304-11
13	50	40.3	29	1	US-08-677-304-11
14	50	40.3	29	2	US-08-436-703B-3
15	50	40.3	29	2	US-08-436-703B-15
16	50	40.3	32	1	US-08-152-488-13
17	50	40.3	32	1	US-08-303-025-15
18	50	40.3	32	1	US-08-677-304-13
19	50	40.3	32	2	US-08-436-703B-2
20	50	40.3	33	1	US-08-303-025-16
21	50	40.3	33	2	US-08-436-703B-4
22	49	39.5	29	1	US-08-152-488-12
23	49	39.5	29	1	US-08-303-025-14
24	49	39.5	29	1	US-08-677-304-12
25	49	39.5	29	2	US-08-436-703B-16
26	49	39.5	928	4	US-09-134-000C-6590
27	47.5	38.3	100	4	US-09-732-210-745

Sequence 308, Appl
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 11042, A
Sequence 36, Appl
Sequence 7464, Ap
Sequence 8019, Ap
Sequence 32, Appl
Sequence 7010, Ap
Sequence 4752, Ap
Sequence 7225, Ap
Sequence 8679, Ap
Sequence 3699, Ap
Sequence 4511, Ap
Sequence 9, Appli
Sequence 3847, Ap

ALIGNMENTS

RESULT.1
US-08-416-711-1
; Sequence 1, Application US/08416711
; Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: BOUHAROUN-TAYOUN, HASNAQ
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; TITLE OF INVENTION: PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P. C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-416-711-1

Query Match 93.5%; Score 116; DB 3; Length 64;
Best Local Similarity 92.0%; Pred. No. 4.3e-10;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
:|||||
Db 1 HERAKNAYOKANQAVLKAKEASSYD 25

RESULT 2
US-09-356-497-1
; Sequence 1, Application US/09356497
; Patent No. 6472519
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-356-497-1

Query Match 93.5%; Score 116; DB 4; Length 64;
Best Local Similarity 92.0%; Pred. No. 4.3e-10;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
:|||||
Db 1 HERAKNAYOKANQAVLKAKEASSYD 25

RESULT 3
US-08-416-711-2
; Sequence 2, Application US/08416711

Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-416-711-2

Query Match 77.4%; Score 96; DB 3; Length 23;
Best Local Similarity 84.0%; Pred. No. 1.2e-07;
Matches 21; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
:|||||
Db 1 HERAKNAYOKANQAVL--KEASSYD 23

RESULT 4
US-09-356-497-2
; Sequence 2, Application US/09356497
; Patent No. 6472519
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA

; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356.497
; FILING DATE: 19-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/416.711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-356-497-2

Query Match 77.4%; Score 96; DB 4; Length 23;
Best Local Similarity 84.0%; Pred. No. 1.2e-07;
Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 1;
QY 1 YEKAKYAYOKANQAVLKAKESASYD 25
Db 1 HERAKYAYKANQAVL--KESASYD 23

RESULT 5
US-08-303-025-12
; Sequence 12, Application US/08303025
; Patent No. 5614494
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 150 West Jefferson, Suite 2500
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226-4415
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy diskette 3.5" 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS v.6.22
; SOFTWARE: WordPerfect 6.1; ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/303,025
; FILING DATE: 08-SEPT-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/06829
; FILING DATE: 14-AUG-1992
; APPLICATION NUMBER: US 08/152,488
; FILING DATE: 12-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REFERENCE/DOCKET NUMBER: 7WK-060548-00231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-496-7622
; TELEFAX: 313-496-8454
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; DOCUMENT NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
US-08-303-025-12

Query Match 40.3%; Score 50; DB 1; Length 28;
Best Local Similarity 52.4%; Pred. No. 0.74;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 2 EKAKYAYOKANQAVLKAKESAS 22
Db 3 KKAACKAAKKAACKAAKKA 23

RESULT 6
US-08-436-703B-1
; Sequence 1, Application US/08436703B
; Patent No. 5919761
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR
; TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
; TITLE OF INVENTION: WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 6601 Woodward Avenue
; STREET: Suite 1525
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6;
; SOFTWARE: ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,703B
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: N/A
; FILING DATE: N/A
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: 7WK-060548-00233

TELECOMMUNICATION INFORMATION:
TELEPHONE: 313-965-1976
TELEFAX: 313-965-1951
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
US-08-436-703B-1

Query Match 40.3%; Score 50; DB 2; Length 28;
Best Local Similarity 52.4%; Pred. No. 0.74;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYOKANQAVLKAKEAS 22
DB 3 KKAKKAARKAKAKAKAKAA 23

RESULT 7

US-08-152-488-10
Sequence 10, Application US/08152488
Patent No. 5534619
GENERAL INFORMATION:
APPLICANT: Wakefield, Thomas W.
APPLICANT: Andrews, Philip C.
TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
TITLE OF INVENTION: ANTICOAGULATION REVERSAL
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Benita J. Rohm, Esq.
STREET: 512 Springfield Avenue
CITY: Cranford
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07016-1811
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6; ASCII (DOS) Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/152,488
FILING DATE: 12-NOV-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08069
FILING DATE: 14-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: RM-7WG
TELEPHONE: 908-276-3344
TELEFAX: 908-276-5543
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A

PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
PUBLICATION INFORMATION:
DOCUMENT NUMBER: PCT/US92/08069
FILING DATE: 14-AUG-1993
US-08-152-488-10

Query Match 40.3%; Score 50; DB 1; Length 29;
Best Local Similarity 52.4%; Pred. No. 0.77;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYOKANQAVLKAKEAS 22
DB 7 KKAKKAARKAKAKAKAKAA 27

RESULT 8

US-08-152-488-11
Sequence 11, Application US/08152488
Patent No. 5534619
GENERAL INFORMATION:
APPLICANT: Wakefield, Thomas W.
APPLICANT: Andrews, Philip C.
APPLICANT: Stanley, James C.
TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
TITLE OF INVENTION: ANTICOAGULATION REVERSAL
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Benita J. Rohm, Esq.
STREET: 512 Springfield Avenue
CITY: Cranford
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07016-1811
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6; ASCII (DOS) Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/152,488
FILING DATE: 12-NOV-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08069
FILING DATE: 14-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: RM-7WG
TELEPHONE: 908-276-3344
TELEFAX: 908-276-5543
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
PUBLICATION INFORMATION:
DOCUMENT NUMBER: PCT/US92/08069
FILING DATE: 14-AUG-1993
US-08-152-488-11

Query Match

40.3%; Score 50; DB 1; Length 29;

Best Local Similarity 52.4%; Pred. No. 0.77; Indels 5; Mismatches 0; Gaps 0;

Qy 2 EKAKYQKQAVLKAKAAS 22
:||||:||||:||||:
Db 7 KKAKAKAKAKAKAKAKAA 27

RESULT 9

US-08-303-025-10
; Sequence 10, Application US/08303025
; Patent No. 5614494
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 150 West Jefferson, Suite 2500
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226-4415
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy diskette 3.5" 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS v.6.22
; SOFTWARE: WordPerfect 6.1; ASCII (DOS) Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/303,025
; FILING DATE: 08-SEPT-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06829
; FILING DATE: 14-AUG-1992
; APPLICATION NUMBER: US 08/152,488
; FILING DATE: 12-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REFERENCE/DOCKET NUMBER: 7WH-060548-00231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-496-7622
; TELEFAX: 313-496-8454
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; DOCUMENT NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993

US-08-303-025-10

Query Match 40.3%; Score 50; DB 1; Length 29;

Best Local Similarity 52.4%; Pred. No. 0.77; Indels 5; Mismatches 0; Gaps 0;

Matches 11; Conservative 5; Mismatches 0; Indels 5; Gaps 0;

Qy 2 EKAKYQKQAVLKAKAAS 22

:||||:||||:||||:

Db 7 KKAKAKAKAKAKAKAKAA 27

RESULT 10

US-08-303-025-11
; Sequence 11, Application US/08303025
; Patent No. 5614494
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 150 West Jefferson, Suite 2500
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226-4415
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy diskette 3.5" 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS v.6.22
; SOFTWARE: WordPerfect 6.1; ASCII (DOS) Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/303,025
; FILING DATE: 08-SEPT-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06829
; FILING DATE: 14-AUG-1992
; APPLICATION NUMBER: US 08/152,488
; FILING DATE: 12-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REFERENCE/DOCKET NUMBER: 7WH-060548-00231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-496-7622
; TELEFAX: 313-496-8454
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; DOCUMENT NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993

US-08-303-025-11

Query Match 40.3%; Score 50; DB 1; Length 29;

Best Local Similarity 52.4%; Pred. No. 0.77; Indels 5; Mismatches 0; Gaps 0;

Matches 11; Conservative 5; Mismatches 0; Indels 5; Gaps 0;

Qy 2 EKAKYQKQAVLKAKAAS 22

:||||:||||:||||:

Db 7 KKAKAKAKAKAKAKAKAA 27

RESULT 11

US-08-303-025-13
; Sequence 13, Application US/08303025
; Patent No. 5614494
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 150 West Jefferson, Suite 2500
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226-4415
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy diskette 3.5" 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS v.6.22
; SOFTWARE: WordPerfect 6.1; ASCII (DOS) Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/303,025
; FILING DATE: 08-SEPT-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06829
; FILING DATE: 14-AUG-1992
; APPLICATION NUMBER: US 08/152,488
; FILING DATE: 12-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REFERENCE/DOCKET NUMBER: 7WH-060548-00231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-496-7622
; TELEFAX: 313-496-8454
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; DOCUMENT NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993

US-08-303-025-13

Query Match 40.3%; Score 50; DB 1; Length 29;

Best Local Similarity 52.4%; Pred. No. 0.77; Indels 5; Mismatches 0; Gaps 0;

Matches 11; Conservative 5; Mismatches 0; Indels 5; Gaps 0;

Qy 2 EKAKYQKQAVLKAKAAS 22

:||||:||||:||||:

Db 7 KKAKAKAKAKAKAKAKAA 27

;/ TITLE OF INVENTION: ANTICOAGULATION REVERSAL
;/ NUMBER OF SEQUENCES: 16
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Benita J. Rohm, Esq.
;/ STREET: 150 West Jefferson, Suite 2500
;/ CITY: Detroit
;/ STATE: Michigan
;/ COUNTRY: United States of America
;/ ZIP: 48226-4415
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy diskette 3.5" 1.44Mb
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: MS-DOS v.6.22
;/ SOFTWARE: WordPerfect 6.1; ASCII (DOS)Text
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/303,025
;/ FILING DATE: 08-SEPT-1994
;/ CLASSIFICATION: 514
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: PCT/US92/06829
;/ FILING DATE: 14-AUG-1992
;/ APPLICATION NUMBER: US 08/152,488
;/ FILING DATE: 12-NOV-1993
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Rohm, Benita J.
;/ REFERENCE/DOCKET NUMBER: 7WH-060548-00231
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: 313-496-7622
;/ TELEFAX: 313-496-8454
;/ INFORMATION FOR SEQ ID NO: 13:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 29 amino acids
;/ TYPE: amino acid
;/ STRANDEDNESS: N/A
;/ TOPOLOGY: N/A
;/ MOLECULE TYPE: peptide
;/ ORIGINAL SOURCE:
;/ ORGANISM: N/A
;/ PUBLICATION INFORMATION:
;/ AUTHORS: N/A
;/ TITLE: N/A
;/ DOCUMENT NUMBER: PCT/US92/08069
;/ FILING DATE: 14-AUG-1993
;/ US-08-303-025-13

Query Match 40.3%; Score 50; DB 1; Length 29;
Best Local Similarity 52.4%; Pred. No. 0.77;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYOKANOAVLKAKKAS 22
DB 7 KKAKKAAKKAKKAAKKAKKAA 27

RESULT 12
US-08-677-304-10
; Sequence 10, Application US/08677304
; Patent No. 5721212
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 512 Springfield Avenue
; CITY: Cranford
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07016-1811

;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: MS-DOS
;/ SOFTWARE: WordPerfect 6; ASCII (DOS)Text
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/677,304
;/ FILING DATE:
;/ CLASSIFICATION: 530
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/152,488
;/ FILING DATE: 12-NOV-1993
;/ APPLICATION NUMBER: PCT/US92/08069
;/ FILING DATE: 14-AUG-1993
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Rohm, Benita J.
;/ REGISTRATION NUMBER: 28,664
;/ REFERENCE/DOCKET NUMBER: RM-7WG
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: 908-276-3344
;/ TELEFAX: 908-276-5543
;/ INFORMATION FOR SEQ ID NO: 10:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 29 amino acids
;/ TYPE: amino acid
;/ STRANDEDNESS: No. 5721212 Relevant
;/ TOPOLOGY: No. 5721212 Relevant
;/ MOLECULE TYPE: peptide
;/ ORIGINAL SOURCE:
;/ ORGANISM: N/A
;/ PUBLICATION INFORMATION:
;/ AUTHORS: N/A
;/ TITLE: N/A
;/ PUBLICATION INFORMATION:
;/ DOCUMENT NUMBER: PCT/US92/08069
;/ FILING DATE: 14-AUG-1993
;/ US-08-677-304-10

Query Match 40.3%; Score 50; DB 1; Length 29;
Best Local Similarity 52.4%; Pred. No. 0.77;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYOKANOAVLKAKKAS 22
DB 7 KKAKKAAKKAKKAAKKAKKAA 27

RESULT 13
US-08-677-304-11
; Sequence 11, Application US/08677304
; Patent No. 5721212
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 512 Springfield Avenue
; CITY: Cranford
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07016-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6; ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,304

;; FILING DATE: 530
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/152,488
;; FILING DATE: 12-NOV-1993
;; APPLICATION NUMBER: PCT/US92/08069
;; FILING DATE: 14-AUG-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Rohm, Benita J.
;; REGISTRATION NUMBER: 28,664
;; REFERENCE/DOCKET NUMBER: RM-7WG
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 908-276-3344
;; TELEFAX: 908-276-5543
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 29 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: No. 5721212 Relevant
;; TOPOLOGY: No. 5721212 Relevant
;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; ORGANISM: N/A
;; PUBLICATION INFORMATION:
;; AUTHORS: N/A
;; TITLE: N/A
;; PUBLICATION INFORMATION:
;; DOCUMENT NUMBER: PCT/US92/08069
;; FILING DATE: 14-AUG-1993
US-08-677-304-11

Query Match 40.3%; Score 50; DB 1; Length 29;
Best Local Similarity 52.4%; Pred. No. 0.77;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYQKANOAVLKAKAAS 22
Db 7 KKAKKAACKAKAKAKAKAA 27

RESULT 14
US-08-436-703B-3
; Sequence 3, Application US/08436703B
; Patent No. 5919761
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR
; TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
; TITLE OF INVENTION: WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 6601 Woodward Avenue
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 6;
; SOFTWARE: ASCII (DOS) text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,703B
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: N/A

;; FILING DATE: N/A
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Rohm, Benita J.
;; REGISTRATION NUMBER: 28,664
;; REFERENCE/DOCKET NUMBER: 7WK-060548-00233
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 313-965-1976
;; TELEFAX: 313-965-1951
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 29 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: N/A
;; TOPOLOGY: N/A
;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; ORGANISM: N/A
;; PUBLICATION INFORMATION:
;; AUTHORS: N/A
;; TITLE: N/A
US-08-436-703B-3

Query Match 40.3%; Score 50; DB 2; Length 29;
Best Local Similarity 52.4%; Pred. No. 0.77;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYQKANOAVLKAKAAS 22
Db 7 KKAKKAACKAKAKAKAKAA 27

RESULT 15
US-08-436-703B-15
; Sequence 15, Application US/08436703B
; Patent No. 5919761
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR
; TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
; TITLE OF INVENTION: WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 6601 Woodward Avenue
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 6;
; SOFTWARE: ASCII (DOS) text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,703B
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: N/A
; FILING DATE: N/A
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: 7WK-060548-00233
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-965-1976
; TELEFAX: 313-965-1951
; INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
US-08-436-703B-15

Query Match 40.3%; Score 50; DB 2; Length 29;
Best Local Similarity 52.4%; Pred. NO. 0.77;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYQKANOAVLKAEAS 22
:|||:|:|:|:|:|:|:|:|:
D 7 KKAKGAKKAKKAKKAKAA 27

Search completed: September 22, 2004, 18:11:00
Job time : 14.7769 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	25	14	US-10-294-770-11
2	124	100.0	25	16	US-10-774-603-11
3	116	93.5	64	14	US-10-294-770-1
4	116	93.5	64	16	US-10-238-741-1
5	116	93.5	64	16	US-10-774-603-1
6	96	77.4	23	14	US-10-294-770-2
7	96	77.4	23	16	US-10-238-741-2
8	96	77.4	23	16	US-10-774-603-2
9	51.5	41.5	866	16	US-10-437-963-203902
10	51.5	41.5	1109	16	US-10-437-963-203905
11	50	40.3	79	12	US-10-393-449-16
12	50	40.3	79	12	US-10-393-449-20
13	50	40.3	79	12	US-10-393-449-66
14	50	40.3	79	12	US-10-393-449-70
15	50	40.3	79	14	US-10-177-725-16
					Sequence 11, Appl
					Sequence 11, Appl
					Sequence 1, Appl
					Sequence 1, Appl
					Sequence 1, Appl
					Sequence 2, Appl
					Sequence 2, Appl
					Sequence 203902,
					Sequence 203905,
					Sequence 16, Appl
					Sequence 20, Appl
					Sequence 66, Appl
					Sequence 70, Appl
					Sequence 16, Appl

```

; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791USODIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-774-602-11

Query Match          100.0%; Score 124; DB 16; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
   |||||
Db 1 YEKAKNAYOKANQAVLKAKEASSYD 25

RESULT 3
US-10-294-770-1
; Sequence 1, Application US/10294770
; Publication No. US20030161840A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759USOCIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-294-770-1

Query Match          93.5%; Score 116; DB 14; Length 64;
Best Local Similarity 92.0%; Pred. No. 2.3e-09;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
   :|||
Db 1 HERAKNAYOKANQAVLKAKEASSYD 25

RESULT 4
US-10-238-741-1
; Sequence 1, Application US/10238741
; Publication No. US20040096466A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE

```

```

; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/238,741
; FILING DATE: 09-Nov-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-238-741-1

Query Match          93.5%; Score 116; DB 16; Length 64;
Best Local Similarity 92.0%; Pred. No. 2.3e-09;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
   :|||
Db 1 HERAKNAYOKANQAVLKAKEASSYD 25

RESULT 5
US-10-774-602-1
; Sequence 1, Application US/10774602
; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791USODIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08

```

;; PRIOR APPLICATION NUMBER: PCT/FR93/01024
;; PRIOR FILING DATE: 1993-10-18
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 1
;; LENGTH: 64
;; TYPE: PRT
;; ORGANISM: Plasmodium falciparum
US-10-774-602-1

Query Match 93.5%; Score 116; DB 16; Length 64;
Best Local Similarity 92.0%; Pred. No. 2.3e-09;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YEKAKNAYQKQANQAVLKAKEASSYD 25
:|:|||||
Db 1 HERAKNAYQKQANQAVLKAKEASSYD 25

RESULT 6
US-10-294-770-2
;; Sequence 2, Application US/10294770
;; Publication No. US20030161840A1
;; GENERAL INFORMATION:
;; APPLICANT: DRUILHE, PIERRE
;; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
;; FILE REFERENCE: 230759USOCIP
;; CURRENT APPLICATION NUMBER: US/10/294,770
;; CURRENT FILING DATE: 2002-11-15
;; PRIOR APPLICATION NUMBER: US 09/356,947
;; PRIOR FILING DATE: 1999-07-19
;; PRIOR APPLICATION NUMBER: US 08/416,711
;; PRIOR FILING DATE: 1995-08-08
;; PRIOR APPLICATION NUMBER: PCT/FR93/01024
;; PRIOR FILING DATE: 1993-10-18
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 2
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Plasmodium falciparum
US-10-294-770-2

Query Match 77.4%; Score 96; DB 14; Length 23;
Best Local Similarity 84.0%; Pred. No. 5.5e-07;
Matches 21; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

Qy 1 YEKAKNAYQKQANQAVLKAKEASSYD 25
:|:|||||
Db 1 HERAKNAYQKQANQAVLKAKEASSYD 23

RESULT 7
US-10-238-741-2
;; Sequence 2, Application US/10238741
;; Publication No. US20040096466A1
;; GENERAL INFORMATION:
;; APPLICANT: DRUILHE, PIERRE
;; BOUHAROUN-TAYOUN, HASNAQ
;; OEUVRAY, CLAUDE
;; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
;; NUMBER OF SEQUENCES: 10
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
;; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
;; CITY: ARLINGTON
;; STATE: VA
;; COUNTRY: USA
;; ZIP: 22202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/238,741
;; FILING DATE: 09-Nov-2002
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/09/356,497
;; FILING DATE: 19-Jul-1999
;; APPLICATION NUMBER: US/08/416,711
;; FILING DATE: 08-AUG-1995
;; APPLICATION NUMBER: PCT/FR93/01024
;; FILING DATE: 18-OCT-1993
;; APPLICATION NUMBER: FR 92/12488
;; FILING DATE: 19-OCT-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: OBLON, NORMAN F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-413-3000
;; TELEFAX: 703-413-2220
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 23 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-238-741-2

Query Match 77.4%; Score 96; DB 16; Length 23;
Best Local Similarity 84.0%; Pred. No. 5.5e-07;
Matches 21; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

Qy 1 YEKAKNAYQKQANQAVLKAKEASSYD 25
:|:|||||
Db 1 HERAKNAYQKQANQAVLKAKEASSYD 23

RESULT 8
US-10-774-602-2
;; Sequence 2, Application US/10774602
;; Publication No. US20040141987A1
;; GENERAL INFORMATION:
;; APPLICANT: DRUILHE, PIERRE
;; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
;; FILE REFERENCE: 248791USODIV
;; CURRENT APPLICATION NUMBER: US/10/774,602
;; CURRENT FILING DATE: 2004-02-10
;; PRIOR APPLICATION NUMBER: US 09/356,947
;; PRIOR FILING DATE: 1999-07-19
;; PRIOR APPLICATION NUMBER: US 10/238,741
;; PRIOR FILING DATE: 2002-09-11
;; PRIOR APPLICATION NUMBER: US 08/416,711
;; PRIOR FILING DATE: 1995-08-08
;; PRIOR APPLICATION NUMBER: PCT/FR93/01024
;; PRIOR FILING DATE: 1993-10-18
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 2
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Plasmodium falciparum
US-10-774-602-2

Query Match 77.4%; Score 96; DB 16; Length 23;
Best Local Similarity 84.0%; Pred. No. 5.5e-07;
Matches 21; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

Qy 1 YEKAKNAYQKQANQAVLKAKEASSYD 25


```

; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (21)-(52)
; OTHER INFORMATION: "Xaa" at positions 21-23, 25-27, 29-30, 32-34, 36-37, 39-41, 43-44
; OTHER INFORMATION: 5, 47-48 and 50-52 can be any amino acid
; US-10-393-449-70

Query Match          40.3%; Score 50; DB 12; Length 79;
Best Local Similarity 63.2%; Pred. No. 9.6;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      3 KAKNAYQKANQAVLKAKEA 21
      ||| ||| : |||||
Db      58 KAKEAEAKAKEAEAKAKEA 76

RESULT 15
US-10-177-725-16
; Sequence 16, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Peele, Beau R.
; TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT
; FILE REFERENCE: A-66900-4/RMS/AMS
; CURRENT APPLICATION NUMBER: US/10/177,725
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 09/415,765
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
; US-10-177-725-16

Query Match          40.3%; Score 50; DB 14; Length 79;
Best Local Similarity 63.2%; Pred. No. 9.6;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      3 KAKNAYQKANQAVLKAKEA 21
      ||| ||| : |||||
Db      58 KAKEAEAKAKEAEAKAKEA 76

Search completed: September 22, 2004, 18:34:03
Job time : 62.9256 secs

```

Search completed: September 22, 2004, 18:34:03
Job time : 62.3256 secs

N;Alternate names: protein YBL0724; protein YBL080c
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: S45428; S45821; S41997; S59225; S41792
R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.
submitted to the EMBL Data Library, May 1994
A;Description: Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces cerevisiae
A;Reference number: S45387
A;Accession: S45428
A;Molecule type: DNA
A;Residues: 1-541 <OBE>
A;Cross-references: EMBL:X79489; NID:g496661; PID:g496701
R;Domdey, H.; Gassenhuber, H.; Obermaier, B.; Piravandi, E.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45816
A;Accession: S45821
A;Molecule type: DNA
A;Residues: 1-541 <DOM>
A;Cross-references: EMBL:Z35841; NID:g536128; PID:g536129; GSPDB:GN00002; MIPS:YBL080c
R;Mulero, J.J.; Rosenthal, J.K.; Fox, T.D.
Genet. 25, 299-304, 1994
A;Title: PET112, a Saccharomyces cerevisiae nuclear gene required to maintain rho(+) mitochondria
A;Reference number: S41997; MUID:94363744; PMID:8082172
A;Accession: S41997
A;Molecule type: DNA
A;Residues: 1-414, 'P', 416-541 <MUL>
A;Cross-references: EMBL:L22072; NID:g347492; PIDN:AAC37508.1; PID:g347493
R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.
Yeast 11, 1103-1112, 1995
A;Title: Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces cerevisiae
A;Reference number: S59184; MUID:96076635; PMID:7502586
A;Accession: S59225
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-541 <OBW>
A;Cross-references: EMBL:X79489; NID:g496661; PIDN:CAA56028.1; PID:g496701
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1994
C;Genetics:
A;Gene: SGD:PET112; MIPS:YBL080c
A;Cross-references: SGD:S0000176; MIPS:YBL080c
A;Map position: 2L
C;Function:
A;Description: involved in mitochondrial gene expression, probably in translation
C;Superfamily: PET112 protein
C;Keywords: mitochondrion; transmembrane protein
P;77-93/Domain: transmembrane #status predicted <TMM>

Query Match 34.0%; Score 50; DB 1; Length 541;
Fast Local Similarity 50.0%; Pred. No. 69;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 4 KKEENMLSHLYVSSKQKE 21
DB 204 KKYQLVRHLHISGGDE 221

Search completed: September 22, 2004, 18:10:01
Job time : 12.6446 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:32 ; Search time 5.78512 Seconds
(without alignments)
252.020 Million cell updates/sec

Title: us-10-774-602-13

Perfect score: 147

Sequence: 1 PEHKEENMLSHLYVSSKDKENISKEND 28

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	55	37.4	1056	1 Z451_MOUSE	Q8C0P7 mus musculus
2	52	35.4	749	1 MADI_YEAST	P40957 saccharomyc
3	52	35.4	993	1 SCPI_MOUSE	Q62209 mus musculus
4	51	34.7	492	1 YE01_YEAST	P40034 saccharomyc
5	51	34.7	572	1 MAOX_MOUSE	P06801 mus musculus
6	50	34.0	541	1 GATB_YEAST	P33893 saccharomyc
7	49	33.3	776	1 ANR5_HUMAN	Q9NU02 homo sapien
8	49	33.3	847	1 GYRA_BUCBP	Q9SA83 buchnera ap
9	49	33.3	997	1 SCPI_RAT	Q03410 rattus norv
10	49	33.3	1956	1 ATX1_PLAFA	Q04956 plasmodium
11	48	32.7	318	1 FMT_MYCPE	Q8X000 mycoplasma
12	48	32.7	504	1 ATPA_CYACA	Q9TM26 cyanidium c
13	48	32.7	845	1 SCPI_MESAU	Q60563 mesocricetu
14	48	32.7	1061	1 Z451_MOUSE	Q9Y465 homo sapien
15	48	32.7	1679	1 GCC2_MOUSE	Q8CG93 mus musculus
16	48	32.7	3584	1 NBEA_DROME	Q9W462 drosophila
17	47.5	32.3	646	1 SSRP_ARATH	Q05153 arabidopsis
18	47.5	32.3	889	1 TPX2_YEAST	P28594 saccharomyc
19	47.5	32.3	913	1 UB26_HUMAN	Q9BXU7 homo sapien
20	47.5	32.3	1085	1 CUI7_SCHPO	P43339 schizosacch
21	47	32.0	204	1 YFHI_YEAST	P43586 saccharomyc
22	47	32.0	482	1 Y138_METJA	Q57602 methanococc
23	47	32.0	530	1 SYK_METJA	Q57959 methanococc
24	47	32.0	902	1 RNE_BUCAI	P57429 buchnera ap
25	47	32.0	1251	1 YMI1_YEAST	Q04545 saccharomyc
26	46.5	31.6	655	1 YKDA_MYCCA	P45615 mycoplasma
27	46	31.3	190	1 YG31_HABIN	P44279 haemophilus
28	46	31.3	317	1 TNMD_RAT	Q9ESC2 rattus norv
29	46	31.3	354	1 YQIB_CABEL	Q09282 caenorhabdi
30	46	31.3	352	1 YKH7_YEAST	P36081 saccharomyc
31	46	31.3	744	1 NUSC_DAMDI	Q32126 dampiera di
32	46	31.3	794	1 Z148_MOUSE	Q61624 mus musculus
33	46	31.3	794	1 Z148_RAT	Q62806 rattus norv

RESULT 1
Z451_MOUSE
ID Z451_MOUSE STANDARD; PRT; 1056 AA.
AC Q8C0P7;
DT 15-MAR-2004 (Rel. 43; Created)
DT 15-MAR-2004 (Rel. 43; Last sequence update)
DT 15-MAR-2004 (Rel. 43; Last annotation update)
DE Zinc finger protein 451.
GN ZNF451.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

ALIGNMENTS

34	46	31.3	885	1	Y143_HUMAN	Q14156 homo sapien
35	46	31.3	1147	1	TEA1_SCHPO	P87061 schizosacch
36	46	31.3	1835	1	MDN1_GIALA	Q8T5T1 giardia lam
37	45.5	31.0	293	1	RRP4_HUMAN	Q13868 homo sapien
38	45.5	31.0	411	1	HFLK_BUCAP	Q8K914 buchnera ap
39	45.5	31.0	573	1	YHDI_SCHPO	Q9P619 schizosacch
40	45	30.6	170	1	Y019_BORBU	Q1051 borrelia bu
41	45	30.6	216	1	EVGI_MOUSE	Q9D981 mus musculus
42	45	30.6	357	1	CADH_POPE	P31657 populus del
43	45	30.6	361	1	COL9_ARATH	Q22800 arabidopsis
44	45	30.6	422	1	TRD2_METJA	Q58759 methanococc
45	45	30.6	424	1	SKP2_MOUSE	Q9Z023 mus musculus

-1- FUNCTION: May be involved in transcriptional regulation.

-1- SUBCELLULAR LOCATION: Nuclear (potential).

-1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.

-1- SIMILARITY: Contains 12 C2H2-type zinc fingers.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AK030088; BAC26778.1; -.
 MGD; MGI:2138298; A1596398.
 InterPro; IPR007087; Znf_C2H2.
 Pfam; PF00096; zf-C2H2; 7.
 SMART; SM00355; Znf_C2H2; 12.
 PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
 PROSITE; PS0157; ZINC_FINGER_C2H2_2; 5.
 Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 Nuclear protein; Repeat.
 FT ZN_FING 169 195 C2H2-TYPE 1.
 FT ZN_FING 212 234 C2H2-TYPE 2 (DEGENERATE).
 FT ZN_FING 253 277 C2H2-TYPE 3.
 FT ZN_FING 315 338 C2H2-TYPE 4 (ATYPICAL).
 FT ZN_FING 362 385 C2H2-TYPE 5.
 FT ZN_FING 494 517 C2H2-TYPE 6.
 FT ZN_FING 527 550 C2H2-TYPE 7.
 FT ZN_FING 604 629 C2H2-TYPE 8 (ATYPICAL).
 FT ZN_FING 634 657 C2H2-TYPE 9.
 FT ZN_FING 665 688 C2H2-TYPE 10.
 FT ZN_FING 751 774 C2H2-TYPE 11.
 FT ZN_FING 787 810 C2H2-TYPE 12.
 SQ SEQUENCE 1056 AA; 120069 MW; 7D8FBC0B50ECA622 CRC64;

Query Match 37.4%; Score 55; DB 1; Length 1056;

Best Local Similarity 38.5%; Pred. No. 15;
 Matches 10; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 3 HKKEENMLSHLYVSSKDKENISKEND 28

DB 683 HYKHHSDYVFVSEKTKTSIKTEGD 708

RESULT 2

MAD1 YEAST

ID MAD1 YEAST STANDARD; PRT; 749 AA.

AC P40957;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Spindle assembly checkpoint component MAD1 (Mitotic MAD1 protein).

GN MAD1 OR YGL086W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OX Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

NCBI_TaxID=4932;

RN [1] _

RP SEQUENCE FROM N.A.

RC STRAIN=W303;

RX MEDLINE=96042315; PubMed=7593191;

RA Hardwick K.G., Murray A.W.;

RT "Mad1p, a phosphoprotein component of the spindle assembly checkpoint

in budding yeast";

RL J. Cell Biol. 131:709-720 (1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=S288c;

RX MEDLINE=97435481; PubMed=9290212;

RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;

RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae

chromosome VII.";

RL Yeast 13:1077-1090 (1997).

RN [3]

RP INTERACTIONS.

RX MEDLINE=98128031; PubMed=9461437;

RA Hwang L.H., Lau L.F., Smith D.L., Mistrot C.A., Hardwick K.G.,

RA Hwang E.S., Amon A., Murray A.W.;

RT "Budding yeast Cdc20: a target of the spindle checkpoint.";

RL Science 279:1041-1044 (1998).

CC -!- FUNCTION: Central component of the spindle assembly checkpoint.

CC -!- SUBUNIT: The spindle checkpoint complex is composed of MAD1, MAD2

CC and MAD3. It interacts with CDC20.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- PTM: Becomes hyperphosphorylated when wild-type cells are arrested

CC in mitosis.

CC -!- SIMILARITY: SOME, TO S.POMBE SPBC3D5.04C.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; U14632; AAA91620.1; -.

DR EMBL; 272608; CAA96791.1; -.

DR PIR; A57276; A57276.

DR GenOnline; 141134; -.

DR SGD; S0003054; MAD1.

DR GO; GO:0005643; C:nuclear pore; IDA.

DR GO; GO:0005634; C:nucleus; IDA.

DR GO; GO:0007094; P:mitotic spindle checkpoint; IGI.

DR GO; GO:0006913; P:nucleocytoplasmic transport; IGI.

DR InterPro; IPR008672; MAD.

DR Pfam; PF05557; MAD.1.

KW Mitosis; Coiled coil; Nuclear protein; Phosphorylation.

FT DOMAIN 57 221 COILED COIL (POTENTIAL).

FT DOMAIN 253 324 COILED COIL (POTENTIAL).

FT DOMAIN 330 656 COILED COIL (POTENTIAL).

FT DOMAIN 323 329 POLY-ASP.

FT DOMAIN 330 347 POLY-ASN.

FT DOMAIN 355 363 POLY-ASN.

FT DOMAIN 601 604 POLY-DEU.

SQ SEQUENCE 749 AA; 87651 MW; 312DDE1BE241610D CRC64;

Query Match 35.4%; Score 52; DB 1; Length 749;

Best Local Similarity 42.3%; Pred. No. 26;

Matches 11; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 EHKKEENMLSHLYVSSKDKENISKEN 27

DB 283 ENEKLNKLSQLHVLESQVENIQLEN 308

RESULT 3

SCP1 MOUSE

ID SCP1 MOUSE STANDARD; PRT; 993 AA.

AC Q62209; O09205; P70192; Q62329;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Synaptonemal complex protein 1 (SCP-1 protein).

GN SYCP1 OR SCP1.

OS Mus musculus [Mouse].

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CBA; TISSUE=Testis;

RX MEDLINE=96004899; PubMed=7548215;

RA Sage J., Martin L., Cuzin F., Rassoulzadegan M.;

RT "cDNA sequence of the murine synaptonemal complex protein 1 (SCP1).";

RL Biochim. Biophys. Acta 1263:258-260 (1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Swiss; TISSUE=Testis;

RA Kerr S.M., Taggart M.H., Lee M., Cooke H.J.;

RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 1-149 FROM N.A.

RC STRAIN=C57BL/6;

RA Sage J., Li Y., Martin L., Mattei M.-G., Guenet J.-L., Liu J.G.,
 RL Hoog C., Cuzin F., Rassoulzadegan M.;
 RN Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE OF 95-787 FROM N.A.
 RC STRAIN=ICK; TISSUE=Testis;
 RA Tsuchida J., Nishina Y., Nozaki M., Uchida K., Nishimune Y.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Major component of the transverse filaments of
 CC synaptonemal complexes (SCS), formed between homologous
 CC chromosomes during meiotic prophase.
 CC -1- SUBCELLULAR LOCATION: Nuclear. In tripartite segments of
 CC synaptonemal complexes, between lateral elements in the nucleus.
 CC found only where the chromosome cores are synapsed. Its N-terminus
 CC is found towards the centre of the synaptonemal complex while the
 CC C-terminus extends well into the lateral domain of the
 CC synaptonemal complex (By similarity).
 CC -1- DOMAIN: Consists of an alpha-helical stretch of 700 AA residues,
 CC flanked by N- and C-terminal globular domains. The C-terminal
 CC domain has DNA-binding capacity (By similarity).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z39118; CAA86262.1; -;
 DR EMBL; L41069; AAC64514.1; ALT_INIT.
 DR EMBL; U62864; AAC53335.1; -;
 DR EMBL; U62860; AAC53335.1; JOINED.
 DR EMBL; U62861; AAC53335.1; JOINED.
 DR EMBL; U62862; AAC53335.1; JOINED.
 DR EMBL; U62863; AAC53335.1; JOINED.
 DR EMBL; D88539; BAA13639.1; -;
 DR PIR; S49461; S49461.
 DR MGD; MGI:105931; Sycpl.
 DR GO; GO:000795; C:synaptonemal complex; IDA.
 DR InterPro; IPR008827; SCP-1.
 DR Pfam; PF05483; SCP-1; 1.
 KW Nuclear protein; Meiosis; Cell division; Phosphorylation;
 KW DNA-binding; Coiled coil.
 FT DOMAIN 12 97 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 104 815 COILED COIL (POTENTIAL).
 FT DOMAIN 114 117 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 697 700 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 898 901 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 978 986 ARG/LYS-RICH (BASIC).
 FT CONFLICT 527 527 F -> L (IN REF. 2).
 SQ SEQUENCE 993 AA; 115962 MW; 1A4FA790D64FAFE6 CRC64;
 Query Match 35.4%; Score 52; DB 1; Length 993;
 Best Local Similarity 41.4%; Pred. No. 35;
 Matches 12; Conservative 6; Mismatches 7; Indels 4; Gaps 1;
 QY 2 EHKKE-----ENMLSHLYVSSKDKENISKE 26
 DB 253 EYQEVNKNQVSELLQIAEKNRMKD 281
 RESULT 4
 YEOL_YEAST
 ID YEOL_YEAST STANDARD; PRT; 492 AA.
 AC P40034;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical 56.5 kDa protein in CAJ1-HOM3 intergenic region.
 GN YER051W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX MEDLINE=97313264; PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
 RA Araujo R., Aviles E., Berne A., Brennan T., Carpenter J., Chen E.,
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
 RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
 RA Lin D., Mossdale D., Nakahara K., Namath A., Norgren R., Oefner P.,
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
 RL Nature 387:78-81(1997).
 CC -1- SIMILARITY: TO C.ELEGANS HYPOTHETICAL PROTEIN T26A5.5.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U18796; RAB64586.1; -;
 DR PIR; S50554; S50554.
 DR GerMOnline; 139131; -;
 DR SGD; S0000853; YER051W.
 DR InterPro; IPR007113; Cupin sup.
 DR InterPro; IPR003347; TF_JmJC.
 DR InterPro; IPR001965; Znf_PHD.
 DR Pfam; PF02373; jmjC; 1.
 DR Pfam; PF00628; PHD; 1.
 DR SMART; SM00558; JmjC; 1.
 DR SMART; SM00249; PHD; 1.
 DR KW Hypothetical protein.
 SQ SEQUENCE 492 AA; 56530 MW; 5C8D09352131692B CRC64;
 Query Match 34.7%; Score 51; DB 1; Length 492;
 Best Local Similarity 33.3%; Pred. No. 22;
 Matches 9; Conservative 7; Mismatches 11; Indels 0; Gaps 0;
 QY 1 PEHKKEENMLSHLYVSSKDKENISKEN 27
 DB 67 PNKEGEYLTAVALITQKGRQRNKEN 93
 RESULT 5
 ID MAOX_MOUSE STANDARD; PRT; 572 AA.
 AC P06801;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NADP-dependent malic enzyme (EC 1.1.1.40) (NADP-ME) (Malic enzyme 1).
 GN ME1 OR MOD1 OR MOD-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87109297; PubMed=3805042;
 RA Bagchi S., Wise L.S., Brown M.L., Bregman D., Sul H.S., Rubin C.S.;
 RT "Structure and expression of murine malic enzyme mRNA.
 RT Differentiation-dependent accumulation of two forms of malic enzyme
 RT mRNA in 3T3-L1 cells.";
 RL J. Biol. Chem. 262:1558-1565(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=87098422; PubMed=3541755;

```

RA Bagchi S., Wise L.S., Brown M.L., Sul H.S., Bregman D.B.,
RA Rubin C.S.;
RA "Regulation and structure of murine malic enzyme mRNA.";
RL Ann. N.Y. Acad. Sci. 478:177-92(1986).
CC -!- CATALYTIC ACTIVITY: (S)-malate + NADP(+) = pyruvate + CO(2) +
CC NADPH.
CC
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the malic enzymes family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J02652; AAA39727.1; -
CC EMBL; M26756; AAA39489.1; -
CC FIR; A26683; DEMSNX.
CC SWISS-2DPAGE; P06801; MOUSE.
CC MGI; 97043; Mod1.
CC InterPro; IPR001891; Malic_oxred.
CC Pfam; PF00390; malic; 1.
CC Pfam; PF03949; malic_N; 1.
CC PRINTS; PR00072; MALOXRDFASE.
CC PROSITE; PS00331; MALIC_ENZYMES; 1.
KW Oxidoreductase; NADP.
FT NP BIND 301 NADP (BY SIMILARITY).
FT SEQUENCE 572 AA; 63998 MW; 86B1761D04C2B51E CRC64;
SQ
Query Match 34.7%; Score 51; DB 1; Length 572;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 10 LSHLYVSSKDKENISKEN 27
DB 309 IAHLVVNAMEKGLSKEN 326
RESULT 6
GATB_YEAST STANDARD; PRT; 541 AA.
AC P33893;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable glutamyl-tRNA(Gln) amidotransferase subunit B, mitochondrial
DE precursor (EC 6.3.5.-) (GLU-ADT subunit B) (Cytochrome oxidase
DE assembly factor PET112).
GN PET112 OR YBL080C OR YBL0724.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94363744; PubMed=8082172;
RA Mulero J.J., Rosenthal J.K., Fox T.D.;
RT "PET112, a Saccharomyces cerevisiae nuclear gene required to maintain
RT rho+ mitochondrial DNA.";
RL Curr. Genet. 25:299-304(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=96076635; PubMed=7502586;
RA Obermaier B., Gassenhuber J., Piravandi E., Domdey H.;
RT "Sequence analysis of a 78.6 kb segment of the left end of
RT Saccharomyces cerevisiae chromosome II.";
RL Yeast 11:1103-1112(1995).
CC -!- FUNCTION: Furnishes a means for formation of correctly charged
CC Gln-tRNA(Gln) through the transamidation of misacylated Glu-

```

```

CC tRNA(Gln) in the mitochondria. The reaction takes place in the
CC presence of glutamine and ATP through an activated gamma-phospho-
CC Glu-tRNA(Gln) (Potential).
CC -!- SUBCELLULAR LOCATION: Mitochondrial (Probable).
CC -!- SIMILARITY: Belongs to the gatB/gate family. GatB subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X79489; CAA56028.1; -
CC EMBL; Z35841; CAA84901.1; -
CC EMBL; L22072; AAC37508.1; -
CC FIR; S45428; S45428.
CC Germonline; 138515; -
CC SGD; S0000176; PET112.
CC GO; GO:0005739; C:mitochondrion; IGI.
CC GO; GO:0009060; P:aerobic respiration; IMP.
CC GO; GO:0007005; P:mitochondrion organization and biogenesis; IMP.
CC GO; GO:0006412; P:protein biosynthesis; IMP.
CC InterPro; IPR004413; GatB_cent.
CC InterPro; IPR006107; GatB_cent.
CC InterPro; IPR006075; GatB_N.
CC InterPro; IPR003789; GatB_Yqey.
CC Pfam; PF01162; GatB; 1.
CC Pfam; PF02934; GatB_N; 1.
CC Pfam; PF02637; GatB_Yqey; 2.
CC TIGRfams; TIGR00133; gatB; 1.
CC PROSITE; PS01234; GATB; 1.
KW Protein biosynthesis; Ligase; Mitochondrion; Transit peptide.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? 541 PROBABLE GLUTAMYL-TRNA(GLN)
FT AMIDOTRANSFERASE SUBUNIT B.
FT CONFLICT 415 415 A -> P (IN REF. 1).
FT SEQUENCE 541 AA; 61842 MW; EE96E82F0F82BBDC CRC64;
Query Match 34.0%; Score 50; DB 1; Length 541;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 4 KKEENMLSHLYVSSKDKKE 21
DB 204 KKYQNLVRHLHSGLDLE 221
RESULT 7
ANR5_HUMAN STANDARD; PRT; 776 AA.
ID ANR5_HUMAN
AC Q9NU02; Q9H6Y9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ankyrin repeat domain protein 5.
GN ANKRD5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark A.G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,

```

```

DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA gyrase subunit A (EC 5.99.1.3).
GN GYRA OR BBP169.
OS Buchnera aphidicola (subsp. Baizongia pistaciae).
OC Bacteria, Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OC NCBI_taxid=135842;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426901; PubMed=12522265;
RY Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
RA Bastolla U., Fernandez J.M., Jinenz L., Postigo M., Silva P.J., A.;
RA Tamames J., Viguera E., Iaborre A., Valencia A., Moran F., Moya A.;
RT "Reductive genome evolution in Buchnera aphidicola.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC -1- FUNCTION: DNA gyrase negatively supercoils closed circular double-
CC stranded DNA in an ATP-dependent manner and also catalyzes the
CC interconversion of other topological isomers of double-stranded
CC DNA rings, including catenanes and knotted rings.
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -1- SUBUNIT: Made up of two chains. The A chain is responsible for DNA
CC breakage and rejoining; the B chain catalyzes ATP hydrolysis. The
CC enzyme forms an A2B2 tetramer.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC -----
DR EMBL: AE014016; AA026902.1; -
DR InterPro: IPR006691; DNA_gyraseA_C.
DR InterPro: IPR002205; DNA_topoisolv.
DR Pfam: PF003989; DNA_gyraseA_C; 6.
DR Pfam: PF00521; DNA_topoisolv; 1.
DR Topoisomerase; Isomerase; DNA-binding; Complete proteome.
KW ACT_SITE 122 122 DNA_CLEAVAGE (BY SIMILARITY).
FT SEQUENCE 847 AA; 95829 MW; BCF78EB59DD939 CRC64;
SQ
Query Match 33.3%; Score 49; DB 1; Length 847;
Best Local Similarity 40.7%; Pred.No. 74;
Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
QY 2 EHKKEENMLSHLYVSKDKENISKEND 28
Db 470 EHKKLISEYKQLFKTSNNLENILKNNN 496
||||| : | : | : | : | :
| : | : | : | : | : | :
RESULT 9
SCPI_RAT
ID SCPI_RAT STANDARD; PRT; 997 AA.
AC Q03410;
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Synaptonemal complex protein 1 (SCP-1 protein).
GN SYCP1 OR SCPI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=101116;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Testis;
RY MEDLINE=93099884; PubMed=1464329;
RA Meuwissen R.L.J., Offenbergh H.H., Dietrich A.J., Rieswijk A.,
RA Iersel M., Heyting C.;
RT "A coiled-coil related protein specific for synapsed regions of
RT meiotic prophase chromosomes."

```

```
RL EMBO J. 11:5091-5100(1992).
CC -!- FUNCTION: Major component of the transverse filaments of
CC synaptonemal complexes (SCS), formed between homologous
CC chromosomes during meiotic prophase.
CC -!- SUBCELLULAR LOCATION: Nuclear. In tripartite segments of
CC synaptonemal complexes, between lateral elements in the nucleus.
CC Found only where the chromosome cores are synapsed. Its N-terminus
CC is found towards the centre of the synaptonemal complex while the
CC C-terminus extends well into the lateral domain of the
CC synaptonemal complex.
CC -!- TISSUE SPECIFICITY: Testis.
CC -!- DEVELOPMENTAL STAGE: Expressed exclusively in meiotic prophase
CC cells.
CC -!- DOMAIN: Consists of an alpha-helical stretch of 700 AA residues,
CC flanked by N- and C-terminal globular domains. The C-terminal
CC domain has DNA-binding capacity.
CC -!- CAUTION: This is a conceptual translation; a frameshift was
CC corrected in position 6 to maximize the similarity with the
CC other species SYCP1 sequences.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X67805; CAA48006.1; ALT_FRAME.
DR InterPro; IPR008827; SCP-1.
DR Pfam; PF05483; SCP-1; 1.
KW Nuclear protein; Meiosis; Cell division; Phosphorylation;
KW DNA-binding; Coiled coil.
FT DOMAIN 108 819 COILED COIL (POTENTIAL).
FT DOMAIN 118 121 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 701 704 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 902 905 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 982 990 ARG/LYS-RICH (BASIC).
FT SEQUENCE 997 AA; 116511 MW; 229D59823FD684BE CRC64;
SQ
Query Match 33.3%; Score 49; DB 1; Length 997;
Best Local Similarity 41.4%; Pred. No. 89;
Matches 12; Conservative 6; Mismatches 7; Indels 4; Gaps 1;

QY 2 EHKKE-----ENMLSHLYVSKKENISKE 26
D 257 EYQKEVNNKENQVSLLIQSTENKMKD 285

RESULT 10
ATX1_PLAFA
ID ATX1_PLAFA STANDARD; PRT; 1956 AA.
AC Q04956;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable cation-transporting ATPase 1 (EC 3.6.3.-).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=79/96;
RX MEDLINE=93132070; PubMed=8421054;
RA Krihnha S., Cowan G., Meade J.C., Wells R.A., Stringer J.R.,
RA Robson K.J.;
RT "A family of cation ATPase-like molecules from Plasmodium
RT falciparum";
RL J. Cell Biol. 120:385-398(1993).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type
CC ATPases). Subfamily V.
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X65738; CAA46646.1; -.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR TIGSFAM6; TIGR01494; ATPase_P-type; 7.
DR PROSITE; PS00154; ATPASE_E1-E2; 1.
KW Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.
FT DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 36 58 POTENTIAL.
FT DOMAIN 59 61 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 62 80 POTENTIAL.
FT DOMAIN 81 407 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 408 427 POTENTIAL.
FT DOMAIN 428 440 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 441 462 POTENTIAL.
FT DOMAIN 463 1818 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1819 1837 POTENTIAL.
FT DOMAIN 1838 1845 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1846 1863 POTENTIAL.
FT DOMAIN 1864 1881 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1882 1905 POTENTIAL.
FT DOMAIN 1906 1928 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1929 1952 POTENTIAL.
FT DOMAIN 1953 1956 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 496 496 PHOSPHORYLATION (BY SIMILARITY).
FT METAL 1760 1760 MAGNESIUM (BY SIMILARITY).
FT METAL 1764 1764 MAGNESIUM (BY SIMILARITY).
FT DOMAIN 246 251 POLY-ASN.
FT DOMAIN 252 256 POLY-LYS.
FT DOMAIN 937 941 POLY-ASN.
FT DOMAIN 1344 1347 POLY-LYS.
FT DOMAIN 1363 1372 POLY-ASN.
FT DOMAIN 1680 1684 POLY-ASN.
SQ SEQUENCE 1956 AA; 230285 MW; AE708AAE99009335 CRC64;

Query Match 33.3%; Score 49; DB 1; Length 1956;
Best Local Similarity 36.1%; Pred. No. 1.9e+02;
Matches 13; Conservative 6; Mismatches 7; Indels 10; Gaps 2;

QY 2 EHKKEENMLS---HL-----YVSKKKNISKEN 27
D 514 QNKQKNMLSDFIHKEMNTESIHSKDNMTIHKN 549

RESULT 11
FMT_MYCPE
ID FMT_MYCPE STANDARD; PRT; 318 AA.
AC Q8EX00;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Methionyl-tRNA formyltransferase (EC 2.1.2.9).
GN FMT OR MYPE500.
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HF-2;
RX MEDLINE=22354719; PubMed=12466555;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
```

```

CC has three main subunits: a, b and c.
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -1- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF022186; AAF13005.1; -.
CC DR InterPro; IPR005294; ATP_synthF1_alph.
CC DR InterPro; IPR000793; ATPase_a/b_C.
CC DR InterPro; IPR000194; ATPase_a/bCentre.
CC DR InterPro; IPR004100; ATPase_a/bNterm.
CC DR InterPro; IPR000790; ATPase_a_C.
CC DR InterPro; IPR009005; F1_ATPase_a/bN.
CC DR Pfam; PF00006; ATP-synt_ab; 1.
CC DR Pfam; PF00306; ATP-synt_ab_C; 1.
CC DR Pfam; PF02874; ATP-synt_ab_N; 1.
CC DR ProDom; PD001099; ATPase_aC; 1.
CC DR TIGRfams; TIGR00962; atpA; 1.
CC DR PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
CC KW ATP synthetase; Chloroplast; Thylakoid; Membrane; CF(1);
CC KM ATP-binding; Hydrolyase; Hydrogen ion transport.
CC FT NP BIND 170 177 ATP (BY SIMILARITY).
CC FT ACT SITE 363 363 BY SIMILARITY.
CC SQ SEQUENCE 504 AA; 54994 MW; EB714398D887B9E0 CRC64;

Query Match 32.7%; Score 48; DB 1; Length 504;
Best Local Similarity 37.9%; Pred.No. 57;
Matches 11; Conservative 6; Mismatches 8; Indels 4; Gaps 1;

QY 4 KXENMLS----HLYVSSKDKENISKEND 28
||| : ||| : : : : :
DB 459 KKLQCLNSNSYPHFYKAIRKESKOLSKENE 487

RESULT 13

```

ID SCPI_MESAU STANDARD; PRI; 845 AA.
 Q050563;
 AC 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Synaptonemal complex protein 1 (SCP-1 protein) (Meiotic chromosome
 DE synapctic protein) (Fragment).
 SCPI OR SYNI.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 NCBI_TaxID=10036;
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Testis;
 RX MEDLINE=95181577; PubMed=7876343;
 RA Dobson M.J., Pearlman R.E., Karaiskakis A., Spyropoulos B.,
 RA Moens P.B.;
 RT "Synaptonemal complex proteins: occurrence, epitope mapping and
 RT chromosome disjunction".
 RL J. Cell Sci. 107:2749-2760(1994).
 CC -I- FUNCTION: Major component of the transverse filaments of
 CC synaptonemal complexes (SCS), formed between homologous
 CC chromosomes during meiotic prophase (By similarity). Has non-
 CC specific DNA binding capability.
 CC -I- SUBCELLULAR LOCATION: Nuclear. In tripartite segments of
 CC synaptonemal complexes, between lateral elements in the nucleus.
 CC Found only where the chromosome cores are synapsed. Its N-terminus
 CC is found towards the centre of the synaptonemal complex while the
 CC C-terminus extends well into the lateral domain of the

```
CC synaptonemal complex.
CC -!- DOMAIN: Consists of an alpha-helical stretch of 700 AA residues,
CC flanked by N- and C-terminal globular domains. The C-terminal
CC domain has DNA-binding capacity (By similarity).
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L32978; AAC2039.1; -
CC PIR; I48176; I48176; SCP-1.
CC InterPro; IPR008827; SCP-1.
CC Pfam; PF05483; SCP-1; 1.
CC NCBI; C12922; SCP-1; 1.
CC DNA-binding; Coiled coil.
CC NON_TER 1
CC DOMAIN 1
CC FT DOMAIN 553 672 COILED COIL (POTENTIAL).
CC FT DOMAIN 753 756 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT DOMAIN 830 838 ARG/LYS-RICH (BASIC).
CC SQ SEQUENCE 845 AA; 99401 MW; D7F28873C824C6A8 CRC64;
CC
CC Query Match 32.7%; Score 48; DB 1; Length 845;
CC Best Local Similarity 41.4%; Pred. No. 1e+02;
CC Matches 12; Conservative 5; Mismatches 8; Indels 4; Gaps 1;
CC
CC QY 2 EHKKE---ENMLSHLYSVSKDKENISKE 26
CC DB 109 EYKKEVNDRENQVSLLIQTEKENKMKD 137
CC
CC RESULT 14
CC Z451.HUMAN
CC ID Z451 HUMAN STANDARD; PRT; 1061 AA.
CC AC Q9Y4E5; Q8N380; Q8TD15; Q9NQW1;
CC DT 15-MAR-2004 (Rel. 43, Created)
CC DT 15-MAR-2004 (Rel. 43, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Zinc finger protein 451 (Coactivator for steroid receptors).
CC GN ZNF451 OR COASTER OR KIA00576.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A. (ISOFORM 2).
CC RA Heldens I.M., Decherling K.J.;
CC RT "Isolation of a novel coactivator for steroid receptors that alters
CC RT the intrinsic activity of the estrogen receptor alpha liganded with
CC RT SERMs.";
CC RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC RN [2]
CC RP SEQUENCE FROM N.A. (ISOFORM 1).
CC RC TISSUE=Brain;
CC RX MEDLINE=98290545; PubMed=9628581;
CC RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
CC RA Nomura N., Ohara O.;
CC RT "Prediction of the coding sequences of unidentified human genes. IX.
CC RT The complete sequences of 100 new cDNA clones from brain which can
CC RT code for large proteins in vitro.";
CC RL DNA Res. 5:31-39(1998).
CC RN [3]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=22935763; PubMed=14574404;
CC RA Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L.,
CC RA Wilming L., Jones M.C., Horton R., Hunt S.E., Scott C.E.,
CC RA Gilbert J.G.R., Clamp M.B., Bethel G., Milne S., Ainscough R.,
CC RA Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S.,
CC RA Babbage A.K., Bagguley C.L., Bailey J., Banerjee R., Barker D.J.,
CC Barlow K.F., Bates K., Beare D.M., Beasley H., Beasley O., Bird C.P.,
CC Blakey S., Bray-Allen S., Brook J., Brown A.J., Brown J.Y.,
CC Burford D.C., Burrill W., Burton J., Carder C., Carter N.P.,
CC Chapman J.C., Clark S.Y., Clark G., Clee C.M., Clegg S., Cobley V.,
CC Collier R.E., Collins J.B., Colman L.K., Corby N.R., Coville G.J.,
CC Culley K.M., Dhami P., Davies J., Dunn M., Earthrowl M.E.,
CC Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A.,
CC Frankland J., French L., Garner P., Garnett J., Ghori M.J.,
CC Gilby L.M., Gillson C.J., Glithero R.J., Grafham D.V., Grant M.,
CC Gribble S., Griffiths C., Griffiths M.N.D., Hall R., Halls K.S.,
CC Hammond S., Harley J.L., Hart E.A., Heath P.D., Heathcote R.,
CC Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Huckle E.,
CC Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A.,
CC Kay M., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C.,
CC Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C.R., Lloyd D.M.,
CC Loveland J.E., Lovell J., Martin S., Mashreghi-Mohammadi M.,
CC Maslen G.L., Matthews L., McCann O.T., McLaren S.J., McLeay K.,
CC McMurray A., Moore M.J.F., Mullikin J.C., Niblett D., Nickerson T.,
CC Novik K.L., Oliver K., Overton-Larty E.K., Parker A., Patel R.,
CC Pearce A.V., Peck A.I., Phillimore B.J.C.T., Phillips S., Plumb R.W.,
CC Porter K.M., Ramsey Y., Randy S.A., Rice C.M., Ross M.T., Searle S.M.,
CC Sehra H.K., Sheridan E., Skuce C.D., Smith S., Smith M., Spraggon L.,
CC Squares S.L., Steward C.A., Sycamore N., Tamlyn-Hall G., Tester J.,
CC Theaker A.J., Thomas D.W., Thorpe A., Tracey A., Tromans A., Tubby B.,
CC Wall M., Wallig J.M., West A.P., White S.S., Whitehead S.L.,
CC Whitaker H., Wild A., Willey D.J., Wilmer T.E., Wood J.M., Wray P.W.,
CC Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A.,
CC Durbin R., Hubbard T., Sulston J.E., Dunham I., Rogers J., Beck S.;
CC RT "The DNA sequence and analysis of human chromosome 6.";
CC RL Nature 425:805-811(2003).
CC RN [4]
CC RP SEQUENCE FROM N.A. (ISOFORM 3).
CC RC TISSUE=Brain;
CC RX MEDLINE=22388257; PubMed=12477932;
CC RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
CC RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
CC RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
CC RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
CC RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
CC RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
CC RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
CC RA Raha S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
CC RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
CC RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
CC RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
CC RA Fahey J., Helton E., Kettman M., Maman A.C., Rodriguez S., Sanchez A.,
CC RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
CC RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
CC RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
CC RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
CC RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
CC RT "Generation and initial analysis of more than 15,000 full-length
CC RT human and mouse cDNA sequences.";
CC RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May be involved in transcriptional regulation.
CC -!- Coactivator for steroid receptors.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9Y4E5-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9Y4E5-2; Sequence=VSP_008624;
CC Name=3;
CC IsoId=Q9Y4E5-3; Sequence=VSP_008625, VSP_008626;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: Contains 11 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```


use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC  EMBL; AY055204; AAL17975.1; -
CC  EMBL; AB011148; BAA25502.1; ALT_INIT.
CC  EMBL; AL363111; CAC01300.1; -
CC  EMBL; BC021712; AAH21712.2; -
CC  PIR; T00341; T00341.
CC  GenBank; HGNC:21091; ZNF451.
CC  InterPro; IPR007087; Znf_C2H2.
CC  Pfam; PF00096; zf-C2H2; 10.
CC  SMART; SMO0355; Znf_C2H2; 10.
CC  PROSITE; PS00028; ZINC_FINGER_C2H2_1; 9.
CC  PROSITE; PS0157; ZINC_FINGER_C2H2_2; 5.
CC  Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
CC  Nuclear protein; Repeat; Alternative splicing.
CC  ZN_FING 169 195
CC  ZN_FING 253 277
CC  ZN_FING 315 337
CC  ZN_FING 362 386
CC  ZN_FING 498 521
CC  ZN_FING 531 554
CC  ZN_FING 606 631
CC  ZN_FING 636 659
CC  ZN_FING 667 690
CC  ZN_FING 753 776
CC  ZN_FING 789 812
CC  VARSPLIC 870 917
CC  Missing (in isoform 2).
CC  /FTID=VSP_008624.
CC  GNTWKPPLNCKIY -> TVNKNKRLKFPPTS (in
CC  isoform 3).
CC  /FTid=VSP_008625.
CC  Missing (in isoform 3).
CC  /FTid=VSP_008626.
CC  I -> T (IN REF. 1).
CC  VARSPLIC 169 169
CC  VARSPLIC 1061 AA; 121484 MW; 8F0446B9FBFF28FA CRC64;
CC  Query Match 32.7%; Score 48; DB 1; Length 1061;
CC  Best Local Similarity 30.8%; Pred. No. 1.3e+02;
CC  Matches 8; Conservative 9; Mismatches 9; Indels 0; Gaps 0;
CC  SQ SEQUENCE 1061 AA; 121484 MW; 8F0446B9FBFF28FA CRC64;
CC  OY 3 HKKEENMLSHLYSSKDKENISKEND 28
CC  Db 685 HYEHHSDIVVFVSEKTESIKTDD 710
CC  RESULT 15
CC  MOUSE
CC  GCC2 MOUSE STANDARD; PRT; 1679 AA.
CC  Q8CHG3; Q8BR44; Q8R2Q5; O9CT45;
CC  DT 10-OCT-2003 (Rel. 42, Created)
CC  DT 10-OCT-2003 (Rel. 42, Last sequence update)
CC  DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC  DE GRIP and coiled-coil domain-containing protein 2 (Golgi coiled coil
CC  protein GCC185).
CC  GN GCC2 OR KIAA0336.
CC  OS Mus musculus (Mouse).
CC  OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC  OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC  OX NCBI_TaxID=10090;
CC  RN [1]
CC  RP SEQUENCE OF 1-769 FROM N.A.
CC  RC TISSUE=Brain;
CC  RX MEDLINE=22398257; PubMed=12477932;
CC  RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
CC  RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
CC  RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
CC  RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haeh F.,
CC  RA Dlatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
CC  RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
CC  RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,

```

```

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
RN [2]
RN SEQUENCE OF 1-442 FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Brain, and Embryo;
RC MEDLINE=22354683; PubMed=12466851;
RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaado I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuoka H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Clothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Varardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirokane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [3]
RN SEQUENCE OF 49-1679 FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=22579291; PubMed=12693553;
RX Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT Prediction of the coding sequences of mouse homologues of KIAA gene:
RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries";
RL DNA Res. 10:35-48(2003).
CC -I- FUNCTION: Probably involved in maintaining Golgi structure (By
CC similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic; peripheral membrane protein
CC associated with the trans-Golgi network (By similarity).
CC -I- DOMAIN: Extended rod-like protein with coiled-coil domains.
CC -I- SIMILARITY: Belongs to the golgin family.
CC -I- SIMILARITY: Contains 1 GRIP domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

DR	ENBL;	BC027339;	AAH27339.1;	ALT_INIT.
DR	ENBL;	MGD;	NGI1917547;	2600014C0IRik.
DR	ENBL;	InterPro;	IPR002317;	GRIP domain.
DR	ENBL;	InterPro;	IPR023017;	Spectrin.
DR	ENBL;	Pfam;	PF01445;	GRIP; 1.
DR	ENBL;	PROSITE;	PS00913;	GRIP; 1.
DR	ENBL;	Golgi stack;	Coiled coil.	
KW	DOMAIN	31	1613	COILED COIL
FT	DOMAIN	83	501	GLU-GRIP
FT	DOMAIN	1604	1654	GRIP.
FT	CONFLICT	440	440	I -> M (
FT	SEQUENCE	1679	AA;	194443 MW;
SQ				652F

Query Match	32.7%	Score 48;	DB 1;	Length 1679;
Local Similarity	36.4%	Pred. No. 2.1e+02;		
Conservative	8;	Mismatches 7;	Indels 0;	Gaps 0;

Qy 3 HKKEENMLSHLYVSSXKXENIS 24
 |::| : || :|| :|| :|| ::
D6 428 HOKEVSELSETFISGSEKEKLA 449

Search completed: September 22, 2004, 18:07:00
Job time : 6.78512 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:02:42 ; Search time 10.2645 Seconds
(without alignments)
253.026 Million cell updates/sec

Title: US-10-774-602-12
Perfect score: 150
Sequence: 1 AKBASSYDYLGNWFGGVPKHEKEN 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	40.7	1238	T34929	hypothetical prote
2	57	38.0	382	1 SYECCS	carbamoyl-phosphat
3	57	38.0	382	2 B85484	carbamoyl-phosphat
4	57	38.0	382	2 C90633	carbamoyl-phosphat
5	55	36.7	479	2 C64655	hypothetical prote
6	55	36.7	479	2 H71942	probable outer mem
7	53	35.3	682	2 T47473	receptor-like prot
8	52.5	35.0	181	2 S78336	conserved hypothet
9	52.5	35.0	602	2 G81195	aspartyl-tRNA synt
10	52.5	35.0	602	2 F81831	aspartate-tRNA lig
11	52	34.7	858	2 T18946	probable phospholi
12	51	34.0	271	2 E70240	pfs protein (pfs)
13	50.5	33.7	588	2 E89945	aspartyl-tRNA synt
14	50	33.3	577	2 A64597	aspartate-tRNA lig
15	50	33.3	579	2 C71916	aspartate-tRNA lig
16	49.5	33.0	168	2 T09333	module specific pr
17	49.5	33.0	2025	2 T21588	hypothetical prote
18	49	32.7	587	2 T48582	auxin-regulated pr
19	49	32.7	613	2 T47483	receptor like prot
20	49	32.7	688	2 T48176	receptor like prot
21	48.5	32.3	420	2 AG1385	B. subtilis YvIB p
22	48	32.0	310	2 C83800	ferrochelatase hem
23	48	32.0	557	2 D83478	gamma-glutamyltran
24	48	32.0	564	2 F71975	methyl-accepting c
25	48	32.0	565	2 G64532	methyl-accepting c
26	48	32.0	567	2 E82956	probable potassium
27	48	32.0	682	2 T48175	receptor like prot
28	47	31.3	416	2 T61744	casein kinase I-ep
29	47	31.3	591	2 AG1264	aspartyl-tRNA synt

aspartyl-tRNA synt
maltodextrin glyco
probable amine oxi
amine oxidase-like
fimbrial usher pro
hypothetical prote
isoleucyl-tRNA syn
hypothetical prote
hypothetical prote
xyloglucan endo-1,
acid phosphatase (
receptor like prot
xylan 1,4-beta-xy
geranylgeranyl pyr
geranyl transrans
nitrilotriacetate

ALIGNMENTS

RESULT 1

T34929
hypothetical protein SC3F9.07 SC3F9.07 - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C/Accession: T34929
R/Seeger, K.J.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1998
A/Reference number: Z21562
A/Accession: T34929
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1238 <SEE>
A/Cross-references: EMBL:AL023862; PIDN:CAAL19630.1; GSPDB:GN00070; SCOEDB:SC3F9.07
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCOEDB:SC3F9.07

Query Match 40.7%; Score 61; DB 2; Length 1238;
Best Local Similarity 44.8%; Pred. No. 1.5;
Matches 13; Conservative 2; Mismatches 8; Indels 6; Gaps 1;

QY 1 AKBASSYDYLGNWFGG-----VPEHK 23
DB 531 AKDADSPDLTYGWDGDKGKGLTTHK 559

RESULT 2

SYECCS
carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5) small chain [validated]
N/Alternate names: carbamoyl-phosphate synthetase glutamine chain
C/Species: Escherichia coli
C/Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 01-Mar-2002
C/Accession: A01128; S40555; H64723
R/Lusty, C.J.
submitted to the Protein Sequence Database, April 1984
A/Reference number: A01128
A/Accession: A01128
A/Molecule type: protein
A/Residues: 1-382 <LUS>
R/Yura, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono, K.; Mizobuchi, T.
submitted to the EMBL Data Library, December 1992
A/Description: Systematic sequencing of the Escherichia coli genome: analysis of the 0-2
A/Reference number: S40551
A/Accession: S40555
A/Molecule type: DNA
A/Residues: 1, 2-382 <YUR>
A/Cross-references: EMBL:D10483
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.


```
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: G81195
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: AB1000; MUID:20175755; PMID:10710307
A;Accession: G81195
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-602 <TET>
A;Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PIDN:AAF40903.1; PID:g7225688
A;Experimental source: serogroup B, strain MC58
C;Genetics:
C;Gene: NMB0466
C;Superfamily: lysine-tRNA ligase

Query Match 35.0%; Score 52.5; DB 2; Length 602;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 4 ASSYDYIL-GWEEGGG 18
| : : : : |
Db 480 ARAYDMVLNGWEIGGG 495

RESULT 10
F81831.
aspartate-tRNA ligase (EC 6.1.1.12) NMA2019 [imported] - Neisseria meningitidis (strain
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
C;Accession: F81831
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morell
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: AB1775; MUID:20222556; PMID:10761919
A;Accession: F81831
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-602 <PAR>
A;Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85238.1; PID:g738064
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
C;Gene: aspS; NMA2019
C;Superfamily: lysine-tRNA ligase
C;Keywords: ligase

Query Match 35.0%; Score 52.5; DB 2; Length 602;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 4 ASSYDYIL-GWEEGGG 18
| : : : : |
Db 480 ARAYDMVLNGWEIGGG 495

RESULT 11
T18946
probable phospholipase activating protein C05C10.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
C;Accession: T18946; T24252
R;Matthews, P.
submitted to the EMBL Data Library, February 1995
A;Reference number: Z19049
A;Accession: T18946
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-858 <WIL>
A;Cross-references: EMBL:Z48178; PIDN:CAA88206.1; GSPDB:GN00020; CESP:C05C10.6
```

```
Best Local Similarity 64.7%; Pred. No. 4.1;
Matches 11; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 8 DYILGWEFGGVPPEHKK 24
| : : : : |
Db 387 DAVSGWVFGGV--HKK 401

RESULT 7
T47473
receptor-like protein kinase - Arabidopsis thaliana
N;Alternate names: protein F18N11.90
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 02-Sep-2000
C;Accession: T47473
R;Jordan, N.; Bangert, S.; Wiedemann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.
submitted to the Protein Sequence Database, February 2000
A;Reference number: Z24467
A;Accession: T47473
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-682 <JOR>
A;Cross-references: EMBL:AL132953
A;Experimental source: cultivar Columbia; BAC clone F18N11
C;Genetics:
A;Map position: 3
A;Note: F18N11.90
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

Query Match 35.3%; Score 53; DB 2; Length 682;
Best Local Similarity 37.1%; Pred. No. 12;
Matches 13; Conservative 0; Mismatches 8; Indels 14; Gaps 1;

QY 6 SYDYLGWEEFGG-----VPEHKKEE 26
| : : : : |
Db 254 SYQYILGWSFSRALLQSLDISKLPVPRPKPE 288

RESULT 8
S78336
conserved hypothetical protein 181 - Odontella sinensis chloroplast
C;Species: chloroplast Odontella sinensis
C;Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 20-Jun-2000
C;Accession: S78336
R;Kowalik, K.V.; Stoebe, B.; Schaffran, P.; Freier, U.
Plant Mol. Biol. Rep. 13, 336-342, 1995
A;Title: The Chloroplast Genome of a Chlorophyll a+c- containing Alga, Odontella sinensis
A;Reference number: S78238
A;Accession: S78336
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-181 <KOW>
A;Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91709.1; PID:g1185226
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
C;Genetics:
C;Gene: ycf4
A;Genome: chloroplast
C;Superfamily: conserved hypothetical protein sl10226
C;Keywords: chloroplast

Query Match 35.0%; Score 52.5; DB 2; Length 181;
Best Local Similarity 50.0%; Pred. No. 3.5;
Matches 11; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 9 YILG---WEFGGVPPEHKKEN 27
| : : : : |
Db 76 YILGTLFDWIDGSGYNEYNKVEN 97

RESULT 9
G81195
aspartyl-tRNA synthetase NMB0466 [imported] - Neisseria meningitidis (strain MC58 serog
C;Species: Neisseria meningitidis
```


This Page Blank (Center)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:32 ; Search time 5.57851 Seconds
(without alignments)
252.020 Million cell updates/sec

Title: US-10-774-602-12
Perfect score: 150
Sequence: 1 AKKASSYDYLWFGGVPHEHKEEN 27

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Number of hits satisfying chosen parameters: 141681

Maximum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	38.0	382	1 CARA_ECOLI	P00907 escherichia
2	56.5	37.7	605	1 SYD_BACHD	Q8Y270 ralsconia s
3	55	36.7	382	1 CARA_ECOL6	Q8f1b1 escherichia
4	52.5	35.0	181	1 YCF4_ODOSI	P49526 odontellia s
5	52.5	35.0	602	1 SYD_NEIMA	Q9jt23 neisseria m
6	52.5	35.0	602	1 SYD_NEIMB	Q9k0u5 neisseria m
7	50.5	33.7	598	1 SYD_STAMM	Q99t19 staphylococ
8	50	33.3	577	1 SYD_HELPY	P56459 helicobacte
9	50	33.3	579	1 SYD_HELPJ	Q9z1l9 helicobacte
10	50	33.3	598	1 SYD_LACPL	Q88vq8 lactobacill
11	48	32.0	310	1 HEMZ_BACHD	Q9kdk9 bacillus ha
12	47.5	31.7	598	1 SYD_STAPP	O8cs99 staphylococ
13	47	31.3	416	1 KCLF_HUMAN	P49674 homo sapien
14	47	31.3	591	1 SYD_LISIN	Q92b74 listeria in
15	47	31.3	591	1 SYD_LISMO	Q8y709 listeria mo
16	47	31.3	835	1 RNFC_PASMU	Q9cnp2 pasteurella
17	46.5	31.0	412	1 PHOA_PENCH	P37274 penicillium
18	46.5	31.0	523	1 C756_CAMME	O04773 campanula m
19	46	30.7	234	1 HISA_STAPP	Q8cq93 staphylococ
20	46	30.7	259	1 KICL_STROG	P29785 streptomyce
21	46	30.7	416	1 KCLF_MOUSE	Q9jmk2 mus musculu
22	46	30.7	439	1 YDDW_ECOLI	P76130 escherichia
23	46	30.7	943	1 SYI_PSEPL	P18330 pseudomonas
24	45.5	30.3	516	1 ACHI_MANSE	P91766 manduca sex
25	45.5	30.3	522	1 SYD_EACSU	Q30208 bacillus su
26	45	30.0	335	1 Y719_ANASP	Q05067 anabaena sp
27	45	30.0	342	1 QUEA_CAMJE	Q9pht9 campylobact
28	45	30.0	359	1 K6P1_ANASP	O8ykg3 anabaena sp
29	45	30.0	537	1 SYE_TREPA	O83679 treponema p
30	45	30.0	598	1 SYD_YERPE	Q8zev0 yersinia pe
31	45	30.0	740	1 RNFC_ECO57	P58324 escherichia
32	45	30.0	740	1 RNFC_ECOLI	P77611 escherichia
33	45	30.0	3712	1 ACVS_CEPAC	P25464 cephalospor

RESULT 1				
CARA_ECOLI	STANDARD;	PRT;	382 AA.	
AC P00907;				
DT 21-JUL-1986 (Rel. 01, Created)				
DT 21-JUL-1986 (Rel. 01, Last sequence update)				
DT 10-OCT-2003 (Rel. 42, Last annotation update)				
DE Carbamoyl-phosphate synthase small chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase glutamine chain).				
GN CARA OR PYRA OR B0032 OR Z0037 OR ECS0035.				
OS Escherichia coli, and				
OS Escherichia coli O157:H7.				
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC Enterobacteriaceae; Escherichia.				
OX NCBI_TaxID=562, 83334;				
RN [1]				
RP SEQUENCE FROM N.A.				
RX MEDLINE=84248072; PubMed=6330744;				
RA Piette J., Nyumoya H., Lusty C.J., Cunin R., Weyens G., Crabbeel M., Charlier D.R.M., Glandsdorff N., Pierard A.;				
RT "DNA sequence of the carA gene and the control region of carAB: tandem promoters, respectively controlled by arginine and the pyrimidines, regulate the synthesis of carbamoyl-phosphate synthetase in Escherichia coli K-12.";				
RT Proc. Natl. Acad. Sci. U.S.A. 81:4134-4138(1984).				
RN [2]				
RP SEQUENCE FROM N.A.				
RX MEDLINE=84248073; PubMed=6377309;				
RA Bouvier J., Pette J.-C., Stragier P.;				
RT "Multiple regulatory signals in the control region of the Escherichia coli carAB operon.";				
RT Proc. Natl. Acad. Sci. U.S.A. 81:4139-4143(1984).				
RN [3]				
RP SEQUENCE FROM N.A.				
RX STRAIN=K12;				
RA MEDLINE=923334977; PubMed=1630901;				
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N., Isono K., Mizobuchi K., Nakata A.;				
RT "Systematic sequencing of the Escherichia coli genome: analysis of the 0-2.4 min region.";				
RT Nucleic Acids Res. 20:3305-3308(1992).				
RN [4]				
RP SEQUENCE FROM N.A.				
RX STRAIN=K12 / MG1655;				
RA MEDLINE=97426617; PubMed=9278503;				
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;				
RT "The complete genome sequence of Escherichia coli K-12.";				
RT Science 277:1453-1474(1997).				
RN [5]				
RP SEQUENCE FROM N.A.				
RX STRAIN=O157:H7 / EDL933 / ATCC 700927;				
RX MEDLINE=21074935; PubMed=11206551;				
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,				

RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocis K.,
RA Apodaca J., Anantharaman I.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RL "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
RL Nature 409:529-533(2001).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Sakagami K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RA "Complete genome sequence of enterohemorrhagic *Escherichia coli*
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [7]
RP SEQUENCE OF 361-382 FROM N.A.
RC MEDLINE=83273669; PubMed=6308632;
RA Nyunoya H., Lusty C.J.;
RA "The carb gene of *Escherichia coli*: a duplicated gene coding for the
RT large subunit of carbamoyl-phosphate synthetase.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:4629-4633(1983).
RN [8]
RP SEQUENCE OF 1-16 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=93096050; PubMed=1334233;
RA Wang M.X., Church G.M.;
RA "A whole genome approach to in vivo DNA-protein interactions in *E.*
RT coli.";
RL Nature 360:606-610(1992).
RN [9]
RP SEQUENCE OF 1-12.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RA "Comparing the predicted and observed properties of proteins encoded
RT in the genome of *Escherichia coli* K-12.";
RL Electrophoresis 18:1259-1313(1997).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=97317071; PubMed=9174345;
RA Thoden J.B., Holden H.M., Wesenberg G., Raushel F.M., Rayment I.;
RA "Structure of carbamoyl phosphate synthetase: a journey of 96 A from
RT substrate to product.";
RL Biochemistry 36:6305-6316(1997).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX MEDLINE=98301363; PubMed=9636022;
RA Thoden J.B., Miran S.G., Phillips J.C., Howard A.J., Raushel F.M.,
RA Holden H.M.;
RA "Carbamoyl phosphate synthetase: caught in the act of glutamine
RT hydrolysis.";
RL Biochemistry 37:8825-8831(1998).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=99190825; PubMed=10089390;
RA Thoden J.B., Raushel F.M., Benning M.M., Rayment I., Holden H.M.;
RA "The structure of carbamoyl phosphate synthetase determined to 2.1-A
RT resolution.";
RL Acta Crystallogr. D 55:8-24(1999).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=99155207; PubMed=10029528;
RA Thoden J.B., Wesenberg G., Raushel F.M., Holden H.M.;
RA "Carbamoyl phosphate synthetase: closure of the B-domain as a result
RT of nucleotide binding.";
RL Biochemistry 38:2347-2357(1999).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=20056035; PubMed=10587438;
RA Thoden J.B., Huang X., Raushel F.M., Holden H.M.;
RA "The small subunit of carbamoyl phosphate synthetase: snapshots along
RT the reaction pathway.";
RL Biochemistry 38:16158-16166(1999).
RN [15]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=99357782; PubMed=10428826;
RA Thoden J.B., Raushel F.M., Wesenberg G., Holden H.M.;
RA "The binding of inosine monophosphate to *Escherichia coli* carbamoyl
RT phosphate synthetase.";
RL J. Biol. Chem. 274:22502-22507(1999).
CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -1- PATHWAY: Arginine biosynthesis.
CC -1- PATHWAY: Pyrimidine biosynthesis; first step.
CC -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain
CC promotes the hydrolysis of glutamine to ammonia, which is used by
CC the large (or ammonia) chain to synthesize carbamoyl phosphate.
CC Tetramer of heterodimers (alpha,beta)4.
CC -1- SIMILARITY: Belongs to the carA family.
CC -1- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10483; BAB96601.1; -.
DR EMBL; AE000113; AAC73143.1; -.
DR EMBL; AE005180; AAG54334.1; -.
DR EMBL; AF002550; BAB33458.1; -.
DR EMBL; AF0017; CAA49615.1; -.
DR PIR; A01128; SYECCS.
DR PIR; B85484; B85484.
DR PIR; C90633; C90633.
DR PDB; 1JDB; 17-JUN-98.
DR PDB; 1A9X; 21-OCT-98.
DR PDB; 1BXR; 20-APR-99.
DR PDB; 1CE8; 26-JUL-99.
DR PDB; 1C30; 10-DEC-99.
DR PDB; 1C30; 10-DEC-99.
DR PDB; 1CS0; 10-DEC-99.
DR PDB; 1KEE; 21-DEC-01.
DR PDB; 1M6V; 13-NOV-02.
DR SWISS-2DPAGE; P00907; COL1.
DR ECO2DBASE; G041.4; 6TH EDITION.
DR EcoGene; EG10134; carA.
DR HAMAP; MF_01209; -; 1.
DR InterPro; IPR006274; CarA_synth_small.
DR InterPro; IPR001317; CP_synthGATase.
DR InterPro; IPR002474; CP_synthsmall.
DR InterPro; IPR00391; GATase 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase_1.
DR PRINTS; PR00099; CPSGATase.
DR PRINTS; PR00096; GATase.
DR TIGRFAMs; TIGR01368; CPSaseIsmall; 1.
DR PROSITE; PS00442; GATASE_TYPE_I; 1.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase;
KW Glutamine amidotransferase; 3D-structure; Complete proteome.
FT DOMAIN 1 189 CFSASE.
FT DOMAIN 190 382 GLUTAMINE AMIDOTRANSFERASE.
FT ACT SITE 269 269 GATASE (BY SIMILARITY).
FT STRAND 4 9
FT TURN 10 11
FT STRAND 14 19
FT STRAND 24 33
FT HELIX 39 43
FT TURN 44 44
FT HELIX 46 48

Query Match 38.0%; Score 57; DB 1; Length 382;
Best Local Similarity 46.7%; Pred. NO. 0.55;
Matches 14; Conservative 4; Mismatches 8; Indels 4; Gaps 2;

QY 1 AKE---ASSYDYILG-WFEGGVPPEHKKEE 26
DB 159 AKVTTAEVSWTQGSWTLTGGLPEAKED 188

RESULT 2

SYD_RALSO
ID SYD_RALSO STANDARD; PRT; 605 AA.
AC Q81270;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aspartyl-tRNA synthetase [EC 6.1.1.12] (Aspartate--tRNA ligase)
DE (AspRS).
GN ASPS OR RSC0466 OR RS04431.

Ralstonia solanacearum (Pseudomonas solanacearum).
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Burkholderiaceae; Ralstonia.
NCBI_TaxID=305;
[1]

SEQUENCE FROM N.A.
STRAIN=GMI1000;
MEDLINE=21681879; PubMed=11823852;
RA Salancubut M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chaudier M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigquier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
RL Nature 415:497-502(2002).

CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; AL646059; CAD13994.1; -;
DR HAMAP; MF 00044; -; 1.
DR InterPro; IPR004524; AppS_bact.
DR InterPro; IPR004115; GAD_dom.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004364; tRNA-synt_2B.
DR InterPro; IPR002312; tRNA-synt_app.
DR InterPro; IPR004365; tRNA_anti.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 2.
DR Pfam; PF01336; tRNA_anti; 1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR TIGRFAMs; TIGR00459; appS_bact; 1.
DR PROSITE; PS00862; AA_tRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.

SQ SEQUENCE 605 AA; 67956 MW; 15EC2C30F7602AF0 CRC64;

Query Match 37.7%; Score 56.5; DB 1; Length 605;
Best Local Similarity 50.0%; Pred. NO. 1.1;
Matches 12; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

QY 4 ASSYDYIL-GWFGGVPPEHKKEE 26
DB 478 AKAYDMVLNGWEIGGSGVRIYREE 501

RESULT 3

CARA_ECOL6
ID CARA_ECOL6 STANDARD; PRT; 382 AA.
AC Q8FLB1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Carbamoyl-phosphate synthase small chain [EC 6.3.5.5] (Carbamoyl-
DE phosphate synthetase glutamine chain).
GN CARA OR C0040.

OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
[1]

SEQUENCE FROM N.A.

STRAIN=O6:H1 / CFT073 / ATCC 700928;
MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + glutamate + carbamoyl phosphate.

CC -1- PATHWAY: Arginine biosynthesis.

CC -1- SUBUNIT: Pyrimidine biosynthesis; first step.

CC promotes the hydrolysis of glutamine to ammonia, which is used by
CC the large (or ammonia) chain to synthesize carbamoyl phosphate (By
CC similarity).

CC -1- SIMILARITY: Belongs to the carA family.

CC -1- SIMILARITY: Contains 1 type-I glutamine amidotransferase domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; AE016755; AAN78538.1; ALT_INIT.
DR HAMAP; MF 01209; -; 1.
DR InterPro; IPR006274; CarA_synth_small.
DR InterPro; IPR001317; CP_synthGATase.
DR InterPro; IPR002474; CP_synthGATase.
DR InterPro; IPR000991; GATase_1.
DR Pfam; PF00988; CPsase_sm_chain; 1.
DR Pfam; PF00117; GATase_1.
DR PRINTS; PR00099; CPsGATase.
DR PRINTS; PR00096; GATase.
DR TIGRFAMs; TIGR01368; CPsaseIIsmall; 1.
DR PROSITE; PS00442; GATASE_TYPE_I; 1.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase;
KW Glutamine biosynthesis; Complete proteome.
FT DOMAIN 1 189
FT DOMAIN 190 382
FT ACT_SITE 269 269
FT ACT_SITE 269 269
SQ SEQUENCE 382 AA; 41415 MW; 308D2BA2EE1478AD CRC64;

Query Match 36.7%; Score 55; DB 1; Length 382;

Best Local Similarity 46.7%; Pred. NO. 1.1;

Matches 14; Conservative 3; Mismatches 9; Indels 4; Gaps 2;

QY 1 AKE---ASSYDYILG-WFEGGVPPEHKKEE 26


```

RX MEDLINE=20175755; PubMed=10710307;
RA Tetelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Chtone H., Clark E.B.,
RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamathavan J.,
RA Gill J., Scariato V., Masignani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MCS8."
RL Science 287:1809-1815 (2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC   diphosphate + L-aspartyl-tRNA(Asp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002403; AAF40903.1; -
CC DR PIR; G81195; G81195.
CC DR HSP; P21889; 1EQR.
CC DR TIGR; NMB0466; -.
CC DR HAMAP; MF_00044; -; 1.
CC DR InterPro; IPR004524; Asps_bact.
CC DR InterPro; IPR004115; GAD_dom.
CC DR InterPro; IPR008994; Nucleic_acid_OB.
CC DR InterPro; IPR004364; tRNA-synt_2.
CC DR InterPro; IPR002312; tRNA-synt_asep.
CC DR InterPro; IPR004365; tRNA_anti.
CC DR InterPro; IPR006195; tRNA_ligase_II.
CC DR Pfam; PF02938; GAD; 1.
CC DR Pfam; PF00152; tRNA-synt_2; 2.
CC DR Pfam; PF01336; tRNA_anti; 1.
CC DR PRINTS; PR01042; TRNASYNTHASP.
CC DR TIGRFAMs; TIGR00459; asps_bact; 1.
CC DR PROSITE; PS00862; AA TRNA_LIGASE II; 1.
CC DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC KW Complete proteome.
CC SQ SEQUENCE 602 AA; 68124 MW; 5DC8A016B0C13E3C CRC64;

Query Match 35.0%; Score 52.5; DB 1; Length 602;
Best Local Similarity 62.5%; Pred. No. 4.3;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 4 ASSYDYIL-GWEFGGG 18
Db 480 ARAYDMVLNGWEIGGG 495

RESULT 7
SYD_STAAM
ID -SYD_STAAM STANDARD; PRT; 588 AA.
AC Q99TL9;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (AspRS).
GN ASPS OR SAV1630 OR SA1456 OR MW1580.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=MU50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Negai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Masumura H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita N., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hirunatu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240 (2001).
CC SEQUENCE FROM N.A.
CC SEQUENCE FROM N.A.
CC SEQUENCE FROM N.A.
CC MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA."
RL Lancet 359:1819-1827 (2002).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC   diphosphate + L-aspartyl-tRNA(Asp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP003363; BAB57792.1; -
CC DR EMBL; AP003134; BAB42722.1; -
CC DR EMBL; AP004827; BAB95445.1; -
CC DR PIR; E89945; E89945.
CC DR HSP; P21889; 1COA.
CC DR HAMAP; MF_00044; -; 1.
CC DR InterPro; IPR004524; Asps_bact.
CC DR InterPro; IPR004115; GAD_dom.
CC DR InterPro; IPR008994; Nucleic_acid_OB.
CC DR InterPro; IPR004364; tRNA-synt_2.
CC DR InterPro; IPR002312; tRNA-synt_asep.
CC DR InterPro; IPR004365; tRNA_anti.
CC DR InterPro; IPR006195; tRNA_ligase_II.
CC DR Pfam; PF02938; GAD; 1.
CC DR Pfam; PF00152; tRNA-synt_2; 2.
CC DR Pfam; PF01336; tRNA_anti; 1.
CC DR PRINTS; PR01042; TRNASYNTHASP.
CC DR TIGRFAMs; TIGR00459; asps_bact; 1.
CC DR PROSITE; PS00862; AA TRNA_LIGASE II; 1.
CC DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC KW Complete proteome.
CC SQ SEQUENCE 588 AA; 66598 MW; 38A6F2C4BF9E9B CRC64;

Query Match 33.7%; Score 50.5; DB 1; Length 588;
Best Local Similarity 58.8%; Pred. No. 8.3;
Matches 10; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Qy 3 EASSYDYIL-GWEFGGG 18
Db 473 EARNAYDIVLNGYELGGG 489

RESULT 8
SYD_HELPY
ID -SYD_HELPY STANDARD; PRT; 577 AA.
AC P56459;

```

```

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (AspRS).
GN ASPS OR HP0617.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Keriavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; AE000576; AAD07682.1; -.
DR PIR; A64597; A64597.
DR HSSP; P36419; LEFW.
DR TIGR; HP0617; -.
DR HAMAP; MF_00044; -.
DR InterPro; IPR004524; Asps_bact.
DR InterPro; IPR004115; GAD_dom.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004364; tRNA-synt_2.
DR InterPro; IPR002312; tRNA-synt_2.
DR InterPro; IPR004365; tRNA-anti.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 2.
DR Pfam; PF01336; tRNA-anti; 1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR TIGRFAMS; TIGR00459; asps_bact; 1.
DR PROSITE; PS50862; AA TRNA LIGASE II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 577 AA; 65601 MW; 29963C376EE9E68F CRC64;

Query Match 33.3%; Score 50; DB 1; Length 577;
Best Local Similarity 52.0%; Pred. No. 9.7;
Matches 13; Conservative 4; Mismatches 6; Indels 2; Gaps 2;

QY 3 EASSYDYL-GWFFGGG-VPEHKE 25
DB 462 EAHAYDVVLNGVELGGGSIHKEE 486

RESULT 9
SYD_HELPJ

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (AspRS).
GN ASPS OR JHP0560.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99120557; PubMed=9923682;
RX Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; AE001489; AAD06143.1; -.
DR PIR; C71916; C71916.
DR HSSP; P36419; LEFW.
DR HAMAP; MF_00044; -.
DR InterPro; IPR004524; Asps_bact.
DR InterPro; IPR004115; GAD_dom.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004364; tRNA-synt_2.
DR InterPro; IPR002312; tRNA-synt_2.
DR InterPro; IPR004365; tRNA-anti.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 2.
DR Pfam; PF01336; tRNA-anti; 1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR TIGRFAMS; TIGR00459; asps_bact; 1.
DR PROSITE; PS50862; AA TRNA LIGASE II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 579 AA; 65616 MW; 4977D97A9562250A CRC64;

Query Match 33.3%; Score 50; DB 1; Length 579;
Best Local Similarity 52.0%; Pred. No. 9.7;
Matches 13; Conservative 4; Mismatches 6; Indels 2; Gaps 2;

QY 3 EASSYDYL-GWFFGGG-VPEHKE 25
DB 462 EAHAYDVVLNGVELGGGSIHKEE 486

RESULT 10
SYD_LACPJ
ID SYD_LACPJ STANDARD; PRT; 598 AA.
AC Q88VQ8;
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last sequence update)

```

```

DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (AsPRS1).
GN ASPS OR LP 1990.
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 8826 / WCFS1;
RX MEDLINE=22480296; PubMed=12566566;
RA Klerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Siezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCFS1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL935257; CAD64363.1; -
CC HAMAP; MF 00044; -; 1.
CC InterPro; IPR004115; GAD dom.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC InterPro; IPR004364; tRNA-synt 2.
CC InterPro; IPR002312; tRNA-synt asp.
CC InterPro; IPR004365; tRNA anti-.
CC InterPro; IPR006195; tRNA_ligase_II.
CC Pfam; PF02938; GAD; 1.
CC Pfam; PF00152; tRNA-synt 2; 1.
CC Pfam; PF01336; tRNA anti; 1.
CC PRINTS; PR01042; TRNASYNTHASP.
CC PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC SEQUENCE 598 AA; 67603 MW; 971B6CDBD645DD6D CRC64;
CC
CC Query Match 33.3%; Score 50; DB 1; Length 598;
CC Best Local Similarity 54.2%; Pred. No. 10;
CC Matches 13; Conservative 3; Mismatches 6; Indels 2; Gaps 2;
CC
CC Qy 4 ASSYDYIL-GWFFGG-VPEHKE 25
CC ||||| : : : : :
CC 475 AQSVDIILNGVLGGSGIRHTRE 498
CC
CC RESULT 11
CC HEMZ BACHD
CC ID HEMZ BACHD STANDARD; PRT; 310 AA.
CC AC Q9KDK9;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Ferrocyclase (EC 4.99.1.1) (Protoheme ferro-lyase) (Heme
CC synthetase).
CC GN HEMH OR BH1203.
CC OS Bacillus halodurans.
CC OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CC OX NCBI_TaxID=86665;
CC RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: Catalyzes the ferrous insertion into protoporphyrin IX.
CC -1- CATALYTIC ACTIVITY: Protoporphyrin + Fe(2+) = protoheme + 2 H(+).
CC -1- PATHWAY: Protoheme biosynthesis; last step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the ferrochelatase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP001511; BAB04922.1; -
CC PIR; C83800; C83800.
CC HSSP; P32396; 1C1H.
CC HAMAP; MF 00323; -; 1.
CC InterPro; IPR001015; Ferrochelatase.
CC Pfam; PF00762; Ferrochelatase; 1.
CC ProDom; PD002792; Ferrochelatase; 1.
CC TIGRFAMs; TIGR00109; hemH; 1.
CC PROSITE; PS00534; FERROCHELATASE; 1.
CC Porphyrin biosynthesis; Heme biosynthesis; Lyase; Iron;
CC Complete proteome.
CC METAL 181 181 IRON (BY SIMILARITY).
CC METAL 262 262 IRON (BY SIMILARITY).
CC SQ SEQUENCE 310 AA; 35271 MW; AA28CE33BAB1FF21 CRC64;
CC
CC Query Match 32.0%; Score 48; DB 1; Length 310;
CC Best Local Similarity 42.9%; Pred. No. 10;
CC Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
CC
CC Qy 1 AKEASSYDYILGWFFGGVPE 21
CC ||||| : : : : :
CC 206 AKGAGITNVAVGWQSGNTPE 226
CC
CC RESULT 12
CC SYD STAEF
CC ID SYD STAEF STANDARD; PRT; 588 AA.
CC AC Q8CS99;
CC DT 15-MAR-2004 (Rel. 43, Created)
CC DT 15-MAR-2004 (Rel. 43, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
CC (AsPRS).
CC DE ASPRS.
CC GN ASPS OR SE1111.
CC OS Staphylococcus epidermidis.
CC OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
CC OX NCBI_TaxID=1282;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX STRAIN=ATCC 12228;
CC PubMed=12950922;
CC RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
CC Qian Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
CC Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
CC "Genome-based analysis of virulence genes in a non-biofilm-forming
CC Staphylococcus epidermidis strain (ATCC 12228).";
CC RL Mol. Microbiol. 49:1577-1593(2003).
CC CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC

```


[illegible]


```

RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a:
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rueniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui A., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Gobel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Roland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).
CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL591979; CAC99597.1; -.
DR PIR; AG1264; AG1264.
DR List18c; LM001519; -.
DR HAMAP; MF_00044; -.
DR InterPro; IPR004524; Asps_bact.
DR InterPro; IPR004115; GAD_dom.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004364; tRNA-synt_2.
DR InterPro; IPR002312; tRNA-synt_2.
DR InterPro; IPR004365; tRNA_anti.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 2.
DR Pfam; PF01336; tRNA_anti_1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR TIGRFAMs; TIGR00459; asps_bact; 1.
DR PROSITE; PS00862; AA_TRNA_LIGASE_II; 1.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
DR Complete proteome.
SQ SEQUENCE 591 AA; 66382 MW; DE143FD009106BBA CRC64;

```

```

Query Match      31.3%; Score 47; DB 1; Length 591;
Best Local Similarity 50.0%; Pred.No. 28;
Matches 12; Conservative 5; Mismatches 5; Indels 2; Gaps 2;

```

```

QY      4 ASSYDYIL-GWEPGGG-VPEHKK 25
DB      473 AEAYDIVLNGYEIGGGSLRIYKKE 496

```

Search completed: September 22, 2004, 18:06:59
Job time : 7.57851 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:32 ; Search time 5.16529 Seconds
(without alignments)
252.020 Million cell updates/sec

Title: US-10-774-602-11
Perfect score: 124
Sequence: 1 YEKANAYQKANOVLKAEASYD 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Maximum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	53	42.7	463	1	K10_DROME
2	50	40.3	106	1	GLPE_VIBCH
3	48	38.7	988	1	TNPS_PSEAE
4	47.5	38.3	100	1	RL23_ECOLI
5	47.5	38.3	472	1	UXAC_OCEIH
6	47.5	38.3	1131	1	RFC1_MOUSE
7	47	37.9	51	1	Y070_RICCN
8	47	37.9	119	1	RL20_NEIMA
9	47	37.9	119	1	RL20_NEIMB
10	47	37.9	541	1	CATA_ASCSU
11	47	37.9	567	1	PGTA_MOUSE
12	46	37.1	119	1	RL20_CHRVO
13	46	37.1	119	1	RL20_NITEU
14	46	37.1	338	1	YQJM_BACSU
15	46	37.1	503	1	AMPA_YERPE
16	45	36.3	124	1	RS16_RHIME
17	45	36.3	431	1	IM44_YEAST
18	45	36.3	632	1	NTP1_SFVKA
19	45	36.3	1086	1	RNC_CABEL
20	45	36.3	1189	1	SMC2_CHICK
21	44.5	35.9	273	1	NADC_HELPY
22	44.5	35.9	503	1	AMPA_ECOLI
23	44.5	35.9	503	1	AMPA_SALTY
24	44.5	35.9	503	1	AMPA_SALTY
25	44.5	35.9	693	1	YD01_SCHPO
26	44	35.5	105	1	GLPE_HAEIN
27	44	35.5	155	1	LEA2_CICAR
28	44	35.5	177	1	LEA1_CICAR
29	44	35.5	199	1	WRBA_BRAJA
30	44	35.5	267	1	IF2A_ARCFU
31	44	35.5	231	1	SNAG_ARATH
32	44	35.5	428	1	PROA_TREPA
33	44	35.5	534	1	HTR2_NATPH

RESULT 1
K10_DROME
ID K10_DROME STANDARD; PRT; 463 AA.
AC P13468; O46075; Q9W505;
DT 01-JAN-1990 (Rel. 13, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-binding protein K10 (female sterile protein K10).
GN FS(1)K10 OR EG:30B8.5 OR CG3218.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Spheroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Prost E., Deryckere F., Roos C., Haenlin M., Pantescio V.,
RA Mohier V.;
RT "Role of the oocyte nucleus in determination of the dorsoventral
RT polarity of Drosophila as revealed by molecular analysis of the K10
RT gene."
RL Genes Dev. 2:891-900(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=20196011; PubMed=10731137;
RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu B.,
RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,
RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
RA Modolell J., Peter A., Schoettler P., Werner M., Mourikoti F.,
RA Beinert N., Dove G., Schaefer U., Jaekle H., Bucheton A.,
RA Callister D.M., Campbell L.A., Darlamsou A., Henderson N.S.,
RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
RA Glover D.M.;
RT "From sequence to chromosome: the tip of the X chromosome of D.
RT melanogaster."
RL Science 287:2220-2222(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Richards S., Ashburner M., Henderson S.N.,
RA George R.A., Lewis S.E., Zhang Q., Chen L.X.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butick H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

Q2696 homo sapien
Q08602 rattus norv
Q09032 saccharomyc
Q07093 drosophila
Q21384 saccharomyc
P33438 drosophila
P37267 saccharomyc
P43814 haemophilus
P17486 caenorhabdi
P45730 populus tri
Q04593 pisum sativ
Q08585 leptospira

ALIGNMENTS

[illegible]

"Structure of the *Escherichia coli* S10 ribosomal protein operon.";
Nucleic Acids Res. 13:4521-4526(1985).
[2]

SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=K12 / MG1655;
MEDLINE=97426617; PubMed=9278503;
Blattner F.R.; Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
"The complete genome sequence of *Escherichia coli* K-12.";
Science 277:1453-1474(1997).
[3]

SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
MEDLINE=22388234; PubMed=12471157;
Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
Moley H.L.T., Donnenberg M.S., Blattner F.R.;
"Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic *Escherichia coli*.";
Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
[4]

SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
MEDLINE=21074935; PubMed=11206551;
Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Welch R.A., Blattner F.R.;
"Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
Nature 409:529-533(2001).
[5]

SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
MEDLINE=21156231; PubMed=11258796;
Havashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Iida T., Takami H., Honda T., Sasakiwa C., Ogasawara N., Yasunaga T.,
Kuhara S., Shibata T., Hattori M., Shinagawa H.;
"Complete genome sequence of enterohemorrhagic *Escherichia coli*
O157:H7 and genomic comparison with a laboratory strain K-12.";
DNA Res. 8:11-22(2001).
[6]

SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=K12;
MEDLINE=80092111; PubMed=391594;
Wittmann-Liebold B., Greuer B.;
"Primary structure of protein L23 from the *Escherichia coli*
ribosome.";
FEBS Lett. 108:69-74(1979).
[7]

SEQUENCE FROM N.A.
SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
MEDLINE=22272406; PubMed=12384590;
Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
Jiang L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
Yu J.;
"Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
through comparison with genomes of *Escherichia coli* K12 and O157.";
Nucleic Acids Res. 30:4432-4441(2002).
[8]

SEQUENCE FROM N.A.
SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
MEDLINE=22590274; PubMed=12704152;
Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
Mau B., Perna N.T., Payne S.M., Runyan-Janecky L.J., Zhou S.,


```

KW Activator; Nuclear protein; Zinc-finger; Alternative splicing.
FT DOMAIN 399 477 BRCT.
FT NP_BIND 635 642 ATP (BY SIMILARITY).
FT ZN_FING 734 751 C2HC-TYPE (POTENTIAL).
FT FT 1108 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 354 528 INTERFERON-STIMULATED-RESPONSE-ELEMENT BINDING REGION.
FT VARSPLIC 614 614 Missing (in isoform 2).
FT CONFLICT 66 66 /FTID=VSP_008444.
FT CONFLICT 187 187 Y -> N (IN REF. 3).
FT CONFLICT 254 254 E -> EPDFCLSLFFGQI (IN REF. 4).
FT CONFLICT 559 559 V -> A (IN REF. 5).
FT CONFLICT 945 945 N -> S (IN REF. 4).
FT CONFLICT 1071 1071 S -> N (IN REF. 1).
FT CONFLICT 1104 1104 T -> A (IN REF. 3).
FT CONFLICT 1131 AA; 125984 MW; A6F4F970A7F9EE94 CRC64;
SQ SEQUENCE 1131 AA; 125984 MW; A6F4F970A7F9EE94 CRC64;

Query Match 38.3%; Score 47.5; DB 1; Length 1131;
Best Local Similarity 36.4%; Pred. No. 67;
Matches 12; Conservative 7; Mismatches 5; Indels 9; Gaps 1;

QY 2 EKAKNYQK-----ANQAVLKAKEASSVD 25
   :|::||:|:||||:|:|||||:|:|:
DB 296 QPKSAHRKEACSSPKASAKLALMKAKESSYN 328

RESULT 7
Y070_RICCN STANDARD; PRT; 51 AA.
ID Y070_RICCN AC Q92J37;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein RC0070.
GN RC0070.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=Mallish 7;
RC MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RL "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RT Science 293:2093-2098(2001).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008574; AL02608.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 51 AA; 5983 MW; 928D5580361EA138 CRC64;

...Query Match 37.9%; Score 47; DB 1; Length 51;
Best Local Similarity 45.0%; Pred. No. 3.9;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 YEKAKNAVOKANQAVLKAKE 20
   |::||:|:|:|:|:|:|:|:|:
DB 10 YRKNRAYKKEFEAWLKCEE 29

RESULT 8
RL20_NEIMA
```

```

ID  RL20_NEIMA  STANDARD;  PRT;  119 AA.
AC  Q9JVA1;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  50S ribosomal protein L20.
GN  RPLT OR NMA0932.
OS  Neisseria meningitidis (serogroup A).
OC  Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC  Neisseriaceae; Neisseria.
OX  NCBI_TaxID=65699;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=22491 / Serogroup A / Serotype 4A;
RX  MEDLINE=2022556; PubMed=10761919;
RA  Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA  Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA  Davies R.M., Davis P., Feltwell T., Hamlin N., Holroyd S.,
RA  Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA  Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA  Whitehead S., Spratt B.G., Barrell B.G.;
RT  "Complete DNA sequence of a serogroup A strain of Neisseria
RT  meningitidis 22491."
RL  Nature 404:502-506(2000).
CC  -!- FUNCTION: This protein binds directly to 23s ribosomal RNA and is
CC  necessary for the in vitro assembly process of the 50s ribosomal
CC  subunit. It is not involved in the protein synthesizing functions
CC  of that subunit (By similarity).
CC  -!- SIMILARITY: Belongs to the L20P family of ribosomal proteins.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AL162754; CAB84204.1; -
CC  PIR; E81939; E81939.
CC  HAMAP; MF_00382; -; 1.
CC  InterPro; IPR005813; Ribosomal_L20.
CC  Pfam; PF00453; Ribosomal_L20; 1.
CC  PRINTS; PR00062; RIBOSOMALL20.
CC  ProDom; PD002389; L20; 1.
CC  TIGRFAMs; TIGR01032; rplT_bact; 1.
CC  PROSITE; PS00937; RIBOSOMAL_L20; 1.
CC  KW  Ribosomal protein; rRNA-binding; Complete proteome.
SQ  SEQUENCE 119 AA; 13680 MW; 9B9859BB5550D650 CRC64;

Query Match 37.9%; Score 47; DB 1; Length 119;
Best Local Similarity 43.5%; Pred. No. 8.8;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 KAKNAYQKNAQVLKAKEASSYD 25
Db : |||:|||||:|:|:|
27 RKKNVYRVAKQAVMKAGQYAYRD 49

RESULT 9
RL20_NEIMB  STANDARD;  PRT;  119 AA.
AC  Q9K093;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  50S ribosomal protein L20.
GN  RPLT OR NMB0723.
OS  Neisseria meningitidis (serogroup B).
OC  Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC  Neisseriaceae; Neisseria.
OX  NCBI_TaxID=491;

```

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=MC58 / Serogroup B;
RX  MEDLINE=20175755; PubMed=10710307;
RA  Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA  Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA  Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA  Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA  Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA  Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA  Gill J., Scariato V., Masiagnani V., Pizzo M., Grandi G., Sun L.,
RA  Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT  "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT  MC58."
RL  Science 287:1809-1815(2000).
CC  -!- FUNCTION: This protein binds directly to 23s ribosomal RNA and is
CC  necessary for the in vitro assembly process of the 50s ribosomal
CC  subunit. It is not involved in the protein synthesizing functions
CC  of that subunit (By similarity).
CC  -!- SIMILARITY: Belongs to the L20P family of ribosomal proteins.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AF002427; RAF41136.1; -
CC  PIR; F81165; F81165.
CC  TIGR; NMB0723; -
CC  HAMAP; MF_00382; -; 1.
CC  InterPro; IPR005813; Ribosomal_L20.
CC  Pfam; PF00453; Ribosomal_L20; 1.
CC  PRINTS; PR00062; RIBOSOMALL20.
CC  ProDom; PD002389; L20; 1.
CC  TIGRFAMs; TIGR01032; rplT_bact; 1.
CC  PROSITE; PS00937; RIBOSOMAL_L20; 1.
CC  KW  Ribosomal protein; rRNA-binding; Complete proteome.
SQ  SEQUENCE 119 AA; 13710 MW; 9B98430B554A6650 CRC64;

Query Match 37.9%; Score 47; DB 1; Length 119;
Best Local Similarity 43.5%; Pred. No. 8.8;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 KAKNAYQKNAQVLKAKEASSYD 25
Db : |||:|||||:|:|:|
27 RKKNVYRVAKQAVMKAGQYAYRD 49

RESULT 10
CATA_ASCSU  STANDARD;  PRT;  541 AA.
AC  P90682;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Catalase (EC 1.11.1.6).
GN  CAT.
OS  Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC  Ascarididae; Ascaris.
OX  NCBI_TaxID=6253;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Eckelt V.H.O.;
RL  Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: Occurs in almost all aerobically respiring organisms and
CC  serves to protect cells from the toxic effects of hydrogen
CC  peroxide.
CC  -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.

```

CC -1- COFACTOR: Heme group (By similarity).
 CC -1- SUBCELLULAR LOCATION: Peroxisomal (By similarity).
 CC -1- SIMILARITY: Belongs to the catalase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; Y10611; CAA71618.1; -;
 DR HSP; P04040; IFAJ.
 DR InterPro; IPR002226; Catalase.
 DR Pfam; PF00199; catalase; 1.
 DR PRINTS; PR00067; CATALASE.
 DR ProDom; PD000510; Catalase; 1.
 DR PROSITE; PS00437; CATALASE_1; 1.
 DR PROSITE; PS00438; CATALASE_2; 1.
 DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
 K Peroxisome.
 FT ACT_SITE 74 74 BY SIMILARITY.
 FT ACT_SITE 147 147 BY SIMILARITY.
 FT METAL 357 357 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 541 AA; 61962 MW; 70890E1AECA170B0 CRC64;
 Query Match 37.9%; Score 47; DB 1; Length 541;
 Best Local Similarity 56.2%; Pred. No. 38;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Oy 5 KNAYQKNAQAVLKAKE 20
 Db 493 RNALQKANEAMQKKE 508
 :||| |||||:|:|
 RESULT 11
 PGTA_MOUSE STANDARD; PRT; 567 AA.
 AC Q9JHK4; Q9JLK2;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Geranylgeranyl transferase type II alpha subunit (EC 2.5.1.60) (Rab
 DE geranylgeranyl transferase alpha subunit) (Rab geranyl-
 DE geranyltransferase alpha subunit) (Rab GG transferase alpha) (Rab
 DE GGTase alpha).
 GN RABGGTA.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND DISEASE.
 RC STRAIN=C57BL/6J-GM/Gm, and C57BL/6J;
 RX MEDLINE=20226072; PubMed=10737774;
 RA Dettler J.C., Zhang Q., Mules E.H., Novack E.K., Mishra V.S., Li W.,
 RA McMurtrie B.B., Tchernev V.T., Wallace M.R., Seabra M.C., Swank R.T.,
 RA Kingmore S.K.;
 RA "Rab geranylgeranyl transferase alpha mutation in the gunmetal mouse
 RT reduces Rab prenylation and platelet synthesis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4144-4149(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=C57BL/6J; TISSUE=Kidney;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayaishizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- FUNCTION: Catalyzes the transfer of a geranyl-geranyl moiety from
 CC geranyl-geranyl pyrophosphate to both cysteines in Rab proteins
 CC with an -XXCC, -XCXC and -CCXX C-terminal, such as RAB1A, RAB3A
 CC and RAB5A respectively (By similarity).
 CC -1- CATALYTIC ACTIVITY: 2 geranylgeranyl diphosphate + protein-
 CC cysteine = 2 S-geranylgeranyl-protein + 2 diphosphate.
 CC -1- ENZYME REGULATION: The enzymatic reaction requires the aid of a
 CC Rab escort protein (also called component A) (By similarity).
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit, collectively
 CC called component B (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9JHK4-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9JHK4-2; Sequence=VSP_009113, VSP_009114;
 CC -1- DISEASE: Defects in RABGGTA are the cause of the gunmetal (gm)
 CC phenotype. Mice homozygous for gm have prolonged bleeding,
 CC thrombocytopenia and reduced platelet alpha- and delta-granule
 CC contents.
 CC -1- SIMILARITY: Belongs to the protein prenyltransferase alpha subunit
 CC family.
 CC -1- SIMILARITY: Contains 5 PPTA repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; AF127654; AAF65918.1; -;
 DR ENBL; AF127655; AAF65919.1; -;
 DR ENBL; AF127656; AAF65920.1; -;
 DR ENBL; AF127658; AAF65921.1; -;
 DR ENBL; AF127659; AAF65922.1; -;
 DR ENBL; AF127660; AAF65923.1; -;
 DR ENBL; AF127662; AAF65924.1; -;
 DR ENBL; AK008625; BAB22240.1; -;
 DR MGI; MGI:1860443; Rabggta.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007092; LRR SDS322.
 DR InterPro; IPR002088; PPTA.
 DR InterPro; IPR008940; Prenyl trans.
 DR InterPro; IPR009087; RabGGT_A_insert.
 DR Pfam; PF00560; LRR; 3.
 DR Pfam; PF01239; PPTA; 5.
 DR PRINTS; PR00019; LEURICHRPT.
 KW Transferase; Prenyltransferase; Repeat; Alternative splicing.
 FT REPEAT 47 80 PPTA 1.
 FT REPEAT 91 124 PPTA 2.
 FT REPEAT 127 160 PPTA 3.
 FT REPEAT 162 195 PPTA 4.
 FT REPEAT 210 243 PPTA 5.
 FT VARSPPLIC 337 339 HOE -> DAV (in isoform 2).
 FT /FTId=VSP_009113.
 FT VARSPPLIC 340 567 Missing (in isoform 2).
 FT /FTId=VSP_009114.
 SQ SEQUENCE 567 AA; 64989 MW; 387DA2DAC12C4C0D CRC64;

Query Match 37.9%; Score 47; DB 1; Length 567;
Best Local Similarity 41.7%; Pred.No. 40;
Matches 10; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 EKAKNAVQKANOAVLKAEASSYD 25
 | : ||||| : ||||| : ||||| :
DB 22 EQKLKYQSATQAVFQREAGELD 45

RESULT 12
RL20_CHRVO STANDARD; PRT: 119 AA.
AC QYN3;
ID AC
AD DT 15-MAR-2004 (Rel. 43, Created)
DT DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE DE 50S ribosomal protein L20.
GG RPLT OR CV1351.

CC Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OX Neisseriaceae; Chromobacterium.
NCBI_TaxId=536;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12472 / DSM 30191;
RX MEDLINE=22862880; PubMed=14500782;
RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
RA Alves-Gomes J.A., Andrade E.M., Aratipe J., de Araujo M.F.F.,
RA Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
RA Batista J.S., Belo A., van den Berg C., Bogio M., Bonatto S.,
RA Bordignon J., Brigido M.M., Brito C.A., Brocchi M., Burty H.A.,
RA Camargo A.C., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
RA Carvalho C.B., Cascado J.C.M., Cavada B.S., Chueire L.M.O.,
RA Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
RA Fantinati P., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Fulan L.R.,
RA Gazzinelli R.T., Gomes E.C., Goncalves P.R., Grangeiro T.B.,
RA Gratnapaglia D., Grisard E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
RA Madeira H.M.P., Manfio G.P., Maranhao A.Q., Martins W.S.,
RA di Mauro S.M.Z., de Medeiros S.R.B., Meissner R.V., Moreira M.A.M.,
RA Nascimento E.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA Ramalho-Neto C.E., Reis A.M.M., Rigó L.U., Rondinelli E.,
RA Santos E.B.P., Santos F.R., Schneider M.P.C., Sevanez H.N.,
RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
RA Souza R.C., Steffens M.B.R., Steinidel M., Teixeira S.R., Urmenyi T.,
RA Vettore A., Wasserm R., Zaha A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
RT remarkable and exploitable bacterial adaptability.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
CC -I- FUNCTION: This protein binds directly to 23S ribosomal RNA and is
CC necessary for the in vitro assembly process of the 50S ribosomal
CC subunit. It is not involved in the protein synthesizing functions
CC of that subunit (By similarity).
CC -I- SIMILARITY: Belongs to the L20P family of ribosomal proteins.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)

DR ENBL; AE016914; AAQ59026.1; --
DR HAMAP; MF_00382; -; 1.
DR PROSITE; PS00937; RIBOSOMAL_L20; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.

```

SQ  SEQUENCE      119 AA;  13626 MW;  2D74991792880CEA CRC64;
Query Match      37.1%;  Score 46;  DB 1;  Length 119;
Best Local Similarity 43.5%;  Pred. No. 12;
Matches 10;  Conservative 5;  Mismatches 8;  Indels 0;  Gaps 0;

Qy  3 KAKNAYOKANQAVLKAKEASSYD 25
Db  27 RRKNVYRIAKQAVNMKAGQAYRD 49

RESULT 13
RL20 NITEU      STANDARD;      PRT;  119 AA.
ID  Q82VW4;
AC  Q82VW4;
DT  15-MAR-2004 (Rel. 43, Created)
DT  15-MAR-2004 (Rel. 43, Last sequence update)
DT  15-MAR-2004 (Rel. 43, Last annotation update)
DE  50S ribosomal protein L20.
DE  GN RPLT OR NE0955.
OS  Nitrosomonas europaea.
OC  Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC  Nitrosomonadaceae; Nitrosomonas.
OX  NCBI_TaxID=915;
[1]
RN  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 19718 / IF0 14298;
RX  MEDLINE=22586410; PubMed=12700255;
RA  Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
RA  Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA  Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RA  "Complete genome sequence of the ammonia-oxidizing bacterium and
RT  obligate chemolithoautotroph Nitrosomonas europaea.";
RL  J. Bacteriol. 185:2759-2773(2003).
CC  -!- FUNCTION: This protein binds directly to 23s ribosomal RNA and is
CC  necessary for the in vitro assembly process of the 50S ribosomal
CC  subunit. It is not involved in the protein synthesizing functions
CC  of that subunit (By similarity).
CC  -!- SIMILARITY: Belongs to the L20P family of ribosomal proteins.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC  or send an email to license@isb-sib.ch).
-----
CC  EMBL; BX321859; CAD94866.1; -.
CC  HAMAP; MF 00382; -.
CC  InterPro; IPR005813; Ribosomal_L20.
CC  InterPro; IPR005812; Ribosomal_L20b/o.
CC  Pfam; PF00453; Ribosomal_L20_1.
CC  PRINTS; PR00062; RIBOSOMALL20.
CC  ProDom; PD002389; L20; 1.
CC  TIGRFAMs; TIGR01032; rplT_bact; 1.
CC  PROSITE; PS00937; RIBOSOMAL_L20; 1.
CC  Ribosomal protein; rRNA-binding; Complete proteome.
CC  KW  SEQUENCE  119 AA;  13724 MW;  779B0490A37D6B4B CRC64;

Query Match      37.1%;  Score 46;  DB 1;  Length 119;
Best Local Similarity 43.5%;  Pred. No. 12;
Matches 10;  Conservative 5;  Mismatches 8;  Indels 0;  Gaps 0;

Qy  3 KAKNAYOKANQAVLKAKEASSYD 25
Db  27 RRKNVYRIAKQAVNMKAGQAYRD 49

RESULT 14
YQJM_BACSU      STANDARD;      PRT;  338 AA.
ID  P54550;
AC  P54550;

```


DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DT Probable NADH-dependent flavin oxidoreductase yqjM (BC 1.-.-.-).
 GN YQM OR BSU23820.
 DE Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RX MEDLINE=97124195; PubMed=8969508;
 RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M., Kobayashi Y.;
 RA "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of the Bacillus subtilis genome containing the skin element and many sporulation genes";
 RL Microbiology 142:3103-3111(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaubner-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Rapoport G., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rev M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter P., Scofione F., Skiguchi J., Sekowska A., Serr S.J., Serror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Takakoshi A., Tanaka T., Terpetra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RL "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis";
 RT Nature 390:249-256(1997).
 CC -1- SIMILARITY: TO THE OYE FAMILY OF NADH-DEPENDENT FLAVIN OXIDOREDUCTASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D84432; BAA12619.1; --
 CC EMBL: Z99116; CAB14314.1; --
 CC PIR: E69964; E69964.
 CC Subtilist; E611742; yqjM.
 CC InterPro: IPR001155; Oxidored FMN.
 CC Pfam: PF00724; Oxidored FMN; I.
 CC Hypothetical protein; Oxidoreductase; NAD; FAD; Flavoprotein;
 CC Complete proteome.
 CC SEQUENCE 338 AA; 37583 MW; BDC52D34236326FE CRC64;
 SQ

Query Match 37.1%; Score 46; DB 1; Length 338;
 Best Local Similarity 50.0%; Pred. No. 33;
 Matches 10; Conservative 8; Indels 0; Gaps 0;
 QY 2 EKAKNAYKQANQAVLKAKA 21
 DB 137 EKVKETVQEFKQAAAKA 156

RESULT 15

AMPA_YERPE STANDARD; PRT; 503 AA.
 ID AMPA_YERPE

AC Q8ZBH3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
 DE (LAP) (Leucyl aminopeptidase).
 DE PEPA OR YPO3441 OR Y0746.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G., Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L., Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M., Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G., Feltwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V., Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
 RL "Genome sequence of Yersinia pestis, the causative agent of plague";
 RL Nature 413:523-527(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KIM5 / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P., Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C., Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V., Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R., Perry R.D.;
 RL "Genome sequence of Yersinia pestis KIM";
 RL J. Bacteriol. 184:4601-4611(2002).
 CC -1- FUNCTION: Presumably involved in the processing and regular turnover of intracellular proteins. Catalyzes the removal of unsubstituted N-terminal amino acids from various peptides (By similarity).
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-Xbb-, in which Xaa is preferably Leu, but may be other amino acids including Pro although not Arg or Lys, and Xbb may be Pro.
 CC -1- COFACTOR: Binds 2 manganese ions per subunit (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family M17.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ414157; CAC92671.1; --
 CC EMBL: AE013676; AAM84333.1; --
 CC PIR: AC0418; AC0418.
 CC MEROPS: M17.003; --
 CC HAMAP: MF_00181; --; 1.
 CC InterPro: IPR000819; Peptidase_M17_C.

DR InterPro; IPR008283; Peptidase M17_N.
DR Pfam; PF00883; Peptidase M17; I.
DR Pfam; PF02789; Peptidase M17_N; 1.
DR PRINTS; PR00481; LAMNOPPTDASE.
DR PROSITE; PS00631; CYTOSOL_AP; 1.
KW Hydrolase; Aminopeptidase; Manganese; Complete proteome.
FT ACT_SITE 282 282 POTENTIAL.
FT ACT_SITE 356 356 POTENTIAL.
FT METAL 270 270 MANGANESE 2 (BY SIMILARITY).
FT METAL 275 275 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 293 293 MANGANESE 2 (BY SIMILARITY).
FT METAL 352 352 MANGANESE 1 (BY SIMILARITY).
FT METAL 354 354 MANGANESE 1 AND 2 (BY SIMILARITY).
SQ SEQUENCE 503 AA; 54796 MW; E3A27876304218F2 CRC64;

Query Match 37.1%; Score 46; DB 1; Length 503;
Best Local Similarity 55.6%; Pred. NO. 49;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

3 KAKNAYQKANKQAVLKAKE 20
| : | | | | | | | | | |
118 KGRNTYWKVRQAVETAKE 135

Db

Search completed: September 22, 2004, 18:06:57
Job time : 14.1653 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:03:03 ; Search time 12.719 Seconds
(without alignments)
109.592 Million cell updates/sec

Title: US-10-774-602-12
Perfect score: 150
Sequence: 1 AKRASSYDYLWFGVGGVPEHKKEEN 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	100.0	27	3	US-08-416-711-3
2	150	100.0	27	4	US-09-356-497-3
3	150	100.0	64	3	US-08-416-711-1
4	150	100.0	64	4	US-09-356-497-1
5	50.5	33.7	597	4	US-09-252-991A-32073
6	49	32.7	490	4	US-09-232-225-41
7	49	32.7	509	4	US-09-232-225-35
8	49	32.7	509	4	US-09-232-225-38
9	48	32.0	306	4	US-09-386-642-53
10	48	32.0	319	4	US-09-386-642-12
11	48	32.0	560	4	US-09-252-991A-22343
12	48	32.0	671	4	US-09-252-991A-19016
13	47.5	31.7	593	4	US-09-134-001C-3592
14	47	31.3	168	4	US-09-543-681A-6335
15	47	31.3	416	3	US-09-100-664A-9
16	47	31.3	416	4	US-09-335-983-9
17	47	31.3	416	4	US-09-553-867A-9
18	47	31.3	416	4	US-09-553-867A-16
19	47	31.3	416	4	US-09-553-867A-18
20	47	31.3	416	4	US-09-472-112-1
21	47	31.3	416	4	US-09-252-991A-18079
22	47	31.3	947	4	US-09-252-991A-21398
23	47	31.3	1390	2	US-08-770-544-2
24	47	31.3	1390	4	US-09-579-259-2
25	47	31.3	1390	4	US-09-650-324A-2
26	46.5	31.0	523	3	US-08-606-505B-67
27	46.5	31.0	523	3	US-09-616-990-67

Sequence 4, Appli
Sequence 4, Appli
Sequence 4654, Ap
Sequence 797, App
Sequence 21, Appl
Sequence 15, Appl
Sequence 18, Appl
Sequence 6086, Ap
Sequence 4524, Ap
Sequence 101, App
Sequence 84, Appl
Sequence 84, Appl
Sequence 5965, Ap
Sequence 9709, Ap
Sequence 23, Appl
Sequence 9477, Ap
Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-416-711-3
; Sequence 3, Application US/08416711
; Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: BOUHAROUN-TAYOUN, HASNAQ
; APPLICANT: OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; TITLE OF INVENTION: PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-416-711-3

US-08-416-711-1
; Sequence 1, Application US/08416711

```

; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-356-497-1

Query Match 100.0%; Score 150; DB 4; Length 64;
Best Local Similarity 100.0%; Pred. No. 7.4e-16;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDIILGWFFGGVPEHKKEEN 27
DB 18 AKEASSYDIILGWFFGGVPEHKKEEN 44

RESULT 5
US-09-252-991A-32073
; Sequence 32073, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32073
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32073

Query Match 33.7%; Score 50.5; DB 4; Length 597;
Best Local Similarity 44.0%; Pred. No. 15;
Matches 11; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 2 KEASSYDIILGW-EFGGVPEHKKE 25
DB 563 KPFNAYKHSIGWGDWGVDPDKLE 587

; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-356-497-1

Query Match 100.0%; Score 150; DB 4; Length 64;
Best Local Similarity 100.0%; Pred. No. 7.4e-16;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDIILGWFFGGVPEHKKEEN 27
DB 18 AKEASSYDIILGWFFGGVPEHKKEEN 44

RESULT 5
US-09-252-991A-32073
; Sequence 32073, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32073
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32073

Query Match 33.7%; Score 50.5; DB 4; Length 597;
Best Local Similarity 44.0%; Pred. No. 15;
Matches 11; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 2 KEASSYDIILGW-EFGGVPEHKKE 25
DB 563 KPFNAYKHSIGWGDWGVDPDKLE 587
```

```

RESULT 6
US-09-292-225-41
; Sequence 41, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 41
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-41

Query Match 32.7%; Score 49; DB 4; Length 490;
Best Local Similarity 43.5%; Pred. No. 20;
Matches 10; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 6 SYDIILGWE--FCGGVPEHKKEE 26
DB 206 TYDHGGWENFVGHNAPLYKRPD 228
```

```

RESULT 7
US-09-292-225-35
; Sequence 35, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 35
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-35

Query Match 32.7%; Score 49; DB 4; Length 509;
Best Local Similarity 43.5%; Pred. No. 21;
Matches 10; Conservative 4; Mismatches 7; Indels 2; Gaps 1;
```

QY 6 SYDYILGWE--FGGVPPEHKKEE 26
 Db 225 TYDHGGWENVFHGNAPLYKRPD 247

RESULT 8
 US-09-292-225-38
 ; Sequence 38, Application US/09292225
 ; Patent No. 6455686
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine A.
 ; APPLICANT: Hunter, Shirley Wu
 ; APPLICANT: Weber, Eric R.
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
 ; FILE REFERENCE: AL-2-C3
 ; CURRENT APPLICATION NUMBER: US/09/292,225
 ; CURRENT FILING DATE: 1999-04-15
 ; PRIOR FILING DATE: 1998-09-02
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR FILING DATE: 1998-04-17
 ; PRIOR FILING DATE: 1998-04-17
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 38
 ; TYPE: PRT
 ; ORGANISM: Dermatophagoides farinae
 US-09-292-225-38

Query Match 32.7%; Score 49; DB 4; Length 509;
 Best Local Similarity 43.5%; Pred. No. 21;
 Matches 10; Conservative 4; Mismatches 7; Indels 7; Gaps 1;

QY 6 SYDYILGWE--FGGVPPEHKKEE 26
 Db 225 TYDHGGWENVFHGNAPLYKRPD 247

RESULT 9
 US-09-386-642-53
 ; Sequence 53, Application US/09386642
 ; Patent No. 6420157
 ; GENERAL INFORMATION:
 ; APPLICANT: Darrow, Andrew
 ; APPLICANT: Qi, Jensen
 ; APPLICANT: Andrade-Gordon, Patricia
 ; TITLE OF INVENTION: Zymogen Activation System
 ; FILE REFERENCE: ORT-1028
 ; CURRENT APPLICATION NUMBER: US/09/386,642
 ; CURRENT FILING DATE: 1999-08-31
 ; NUMBER OF SEQ ID NOS: 60
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 53
 ; LENGTH: 306
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene of
 ; OTHER INFORMATION: human protease F in CFEK2 zymogen vector
 US-09-386-642-53

Query Match 32.0%; Score 48; DB 4; Length 306;
 Best Local Similarity 44.4%; Pred. No. 17;
 Matches 8; Conservative 6; Mismatches 4; Indels 4; Gaps 0;

QY 10 ILGWFEFGGVPPEHKKEEN 27
 Db 12 LLGTTFGCGVPDYKDDDD 29

RESULT 10
 US-09-386-642-12
 ; Sequence 12, Application US/09386642
 ; Patent No. 6420157
 ; GENERAL INFORMATION:
 ; APPLICANT: Darrow, Andrew
 ; APPLICANT: Qi, Jensen
 ; APPLICANT: Andrade-Gordon, Patricia
 ; TITLE OF INVENTION: Zymogen Activation System
 ; FILE REFERENCE: ORT-1028
 ; CURRENT APPLICATION NUMBER: US/09/386,642
 ; CURRENT FILING DATE: 1999-08-31
 ; NUMBER OF SEQ ID NOS: 60
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 12
 ; LENGTH: 319
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
 ; OTHER INFORMATION: with homo sapien serine protease catalytic domain
 US-09-386-642-12

Query Match 32.0%; Score 48; DB 4; Length 319;
 Best Local Similarity 44.4%; Pred. No. 18;
 Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 10 ILGWFEFGGVPPEHKKEEN 27
 Db 12 LLGTTFGCGVPDYKDDDD 29

RESULT 11
 US-09-252-991A-22343
 ; Sequence 22343, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 22343
 ; LENGTH: 560
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-22343

Query Match 32.0%; Score 48; DB 4; Length 560;
 Best Local Similarity 56.2%; Pred. No. 34;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 ASSDYILGWEFGGV 19
 Db 379 AVSNTYTLNWDGSGV 394

RESULT 12
 US-09-252-991A-19016
 ; Sequence 19016, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; OTHER INFORMATION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:09:13 ; Search time 58.2397 Seconds
(without alignments)
148.878 Million cell updates/sec

Title: US-10-774-602-12
Perfect score: 150
Sequence: 1 AKEASSYDYLGWFGGVPPEHKKEN 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Maximum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description	
Result No.	Score	Query Match	Length	ID		
1	150	100.0	27	14	US-10-294-770-3	Sequence 3, Appl
2	150	100.0	27	14	US-10-294-770-12	Sequence 12, Appl
3	150	100.0	27	16	US-10-238-741-3	Sequence 3, Appl
4	150	100.0	27	16	US-10-774-602-3	Sequence 3, Appl
5	150	100.0	27	16	US-10-774-602-12	Sequence 12, Appl
6	150	100.0	64	14	US-10-294-770-1	Sequence 1, Appl
7	150	100.0	64	16	US-10-238-741-1	Sequence 1, Appl
8	150	100.0	64	16	US-10-774-602-1	Sequence 1, Appl
9	58	38.7	596	12	US-10-282-122A-51159	Sequence 51159, A
10	58	38.7	599	12	US-10-282-122A-49798	Sequence 49798, A
11	58	38.7	600	12	US-10-282-122A-48020	Sequence 48020, A
12	58	38.7	600	12	US-10-282-122A-50005	Sequence 50005, A
13	57	38.0	382	12	US-10-282-122A-43296	Sequence 43296, A
14	57	38.0	382	14	US-10-210-115-33	Sequence 33, Appl
15	57	38.0	382	15	US-10-369-493-691	Sequence 691, App

16	55	36.7	247	12	US-10-335-977-4888	Sequence 4888, Ap
17	55	36.7	479	9	US-09-881-752A-14	Sequence 14, Appl
18	55	36.7	479	12	US-10-335-977-4890	Sequence 4890, Ap
19	55	36.7	486	12	US-10-335-977-4891	Sequence 4891, Ap
20	54	36.0	432	16	US-10-437-963-123350	Sequence 123350, A
21	52.5	35.0	86	16	US-10-767-701-51357	Sequence 51357, A
22	52.5	35.0	111	14	US-10-091-300-35	Sequence 35, Appl
23	52.5	35.0	111	14	US-10-091-300-51	Sequence 51, Appl
24	52.5	35.0	594	12	US-10-282-122A-65399	Sequence 65399, A
25	52.5	35.0	602	12	US-10-282-122A-66049	Sequence 66049, A
26	52	34.7	378	15	US-10-369-493-360	Sequence 360, App
27	52	34.7	858	15	US-10-369-493-5536	Sequence 5536, Ap
28	52	34.7	858	15	US-10-369-493-5537	Sequence 5537, Ap
29	51.5	34.3	269	9	US-09-888-721-36	Sequence 36, Appl
30	51.5	34.3	282	9	US-09-888-721-38	Sequence 38, Appl
31	51.5	34.3	287	9	US-09-888-721-40	Sequence 40, Appl
32	51.5	34.3	251	9	US-09-888-721-44	Sequence 44, Appl
33	51.5	34.3	296	9	US-09-888-721-42	Sequence 42, Appl
34	51	34.0	66	16	US-10-437-963-106497	Sequence 106497, A
35	51	34.0	263	14	US-10-156-761-9233	Sequence 9233, Ap
36	51	34.0	404	16	US-10-437-963-103597	Sequence 103597, A
37	51	34.0	605	16	US-10-437-963-139769	Sequence 139769, A
38	50.5	33.7	588	9	US-09-815-242-12553	Sequence 12553, A
39	50.5	33.7	588	9	US-09-815-242-12816	Sequence 12816, A
40	50.5	33.7	588	12	US-10-282-122A-44049	Sequence 44049, A
41	50.5	33.7	590	9	US-09-815-242-5344	Sequence 5344, Ap
42	50	33.3	59	12	US-10-424-599-254448	Sequence 254448, A
43	50	33.3	174	12	US-10-335-977-7134	Sequence 7134, Ap
44	50	33.3	238	10	US-09-880-748-2067	Sequence 2067, Ap
45	50	33.3	238	12	US-10-293-418-2067	Sequence 2067, Ap

ALIGNMENTS

RESULT 1
US-10-294-770-3
; Sequence 3, Application US/10294770
; Publication No. US20030161840A1
; GENERAL INFORMATION:
; APPLICANT: DRULHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759USOCIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR FILING DATE: 2002-11-15
; PRIOR FILING DATE: 1999-07-19
; PRIOR FILING DATE: 1995-08-08
; PRIOR FILING DATE: 1995-08-08
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-294-770-3

Query Match	100.0%	Score 150;	DB 14;	Length 27;
Best Local Similarity	100.0%	Pred. No. 1.9e-14;		
Matches	27;	Conservative	0;	Mismatches 0;
Indels	0;	Gaps	0;	

Qy	1	AKEASSYDYLGWFGGVPPEHKKEN	27
Db	1	AKEASSYDYLGWFGGVPPEHKKEN	27

RESULT 2			
US-10-294-770-12			
; Sequence 12, Application US/10294770			
; Publication No. US20030161840A1			
; GENERAL INFORMATION:			

APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759US0CIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
; US-294-770-12

Query Match 100.0%; Score 150; DB 14; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPHEHKEEN 27
DB 1 AKEASSYDYLGWFGGVPHEHKEEN 27

RESULT 3
US-10-238-741-3
; Sequence 3, Application US/10238741
; Publication No. US20040096466A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JERFPERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/238,741
; FILING DATE: 09-Nov-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-238-741-3

Query Match 100.0%; Score 150; DB 16; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPHEHKEEN 27
DB 1 AKEASSYDYLGWFGGVPHEHKEEN 27

RESULT 4
US-10-774-602-3
; Sequence 3, Application US/10774602
; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791US0DIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; US-10-774-602-3

Query Match 100.0%; Score 150; DB 16; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPHEHKEEN 27
DB 1 AKEASSYDYLGWFGGVPHEHKEEN 27

RESULT 5
US-10-774-602-12
; Sequence 12, Application US/10774602
; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791US0DIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18

```

; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-774-602-12

Query Match      100.0%; Score 150; DB 16; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLGWFGGVPHEKKEEN 27
    |||||
Db 1 AKEASSYDYLGWFGGVPHEKKEEN 27

RESULT 6
US-10-294-770-1
; Sequence 1, Application US/10294770
; Publication No. US20030161840A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759USOCIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR FILING DATE: 1995-08-08
; PRIOR FILING DATE: 1993-10-18
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-294-770-1

Query Match      100.0%; Score 150; DB 14; Length 64;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLGWFGGVPHEKKEEN 27
    |||||
Db 18 AKEASSYDYLGWFGGVPHEKKEEN 44

RESULT 7
US-10-238-741-1
; Sequence 1, Application US/10238741
; Publication No. US20040096466A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/238,741
; FILING DATE: 09-Nov-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-238-741-1

Query Match      100.0%; Score 150; DB 16; Length 64;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLGWFGGVPHEKKEEN 27
    |||||
Db 18 AKEASSYDYLGWFGGVPHEKKEEN 44

RESULT 8
US-10-774-602-1
; Sequence 1, Application US/10774602
; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791USODIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-774-602-1

Query Match      100.0%; Score 150; DB 16; Length 64;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLGWFGGVPHEKKEEN 27
    |||||
Db 18 AKEASSYDYLGWFGGVPHEKKEEN 44

```

Db 18 AKEASSYDYLGWFGGVPHEHKEEN 44
|||||

RESULT 9

US-10-282-122A-51159
; Sequence 51159, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51159
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Bordetella pertussis

US-10-282-122A-51159

US-10-282-122A-49798
; Sequence 49798, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49798
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Burkholderia fungorum

US-10-282-122A-51159

US-10-282-122A-48020
; Sequence 48020, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51159
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Bordetella pertussis

US-10-282-122A-51159

US-10-282-122A-49798

US-10-282-122A-51159

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

US-10-282-122A-48020

```
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48020
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-48020

Query Match      38.7%; Score 58; DB 12; Length 600;
Best Local Similarity 54.2%; Pred. No. 9.9;
Matches 13; Conservative 4; Mismatches 5; Indels 2; Gaps 2;

Qy      4 ASSYDYL-GWEPGGG-VPEHKKE 25
Db      480 AKAYDMVLNGWEIGGSGVRIHREE 503

RESULT 12
US-10-282-122A-50005
; Sequence 50005, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48020
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-48020

Query Match      38.7%; Score 58; DB 12; Length 600;
Best Local Similarity 54.2%; Pred. No. 9.9;
Matches 13; Conservative 4; Mismatches 5; Indels 2; Gaps 2;

Qy      4 ASSYDYL-GWEPGGG-VPEHKKE 25
Db      480 AKAYDMVLNGWEIGGSGVRIHREE 503
```

```
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50005
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Burkholderia mallei
US-10-282-122A-50005

Query Match      38.7%; Score 58; DB 12; Length 600;
Best Local Similarity 54.2%; Pred. No. 9.9;
Matches 13; Conservative 4; Mismatches 5; Indels 2; Gaps 2;

Qy      4 ASSYDYL-GWEPGGG-VPEHKKE 25
Db      480 AKAYDMVLNGWEIGGSGVRIHREE 503

RESULT 13
US-10-282-122A-43296
; Sequence 43296, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43296
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-43296

Query Match      38.0%; Score 57; DB 12; Length 382;
Best Local Similarity 46.7%; Pred. No. 8.6;
```

Matches 14; Conservative 4; Mismatches 8; Indels 4; Gaps 2;
QY 1 AKE---ASSVDYILG-WFEGGVPEHKKEE 26
Db 159 AKEVTTAEAYSWTQGSWTLTGGLPEAKKD 188
Search completed: September 22, 2004, 18:34:04
Job time : 59.2397 secs

RESULT 14
US-10-210-115-33
; Sequence 33, Application US/10210115
; Publication No. US20030129708A1
; GENERAL INFORMATION:
; APPLICANT: Ptitsyn, Leonid R
; APPLICANT: Smirnov, Sergey V
; APPLICANT: Altman, Irina B
; APPLICANT: No. US20030129708A1ikova, Anna E
; APPLICANT: Kotliarova, Veronika A
; APPLICANT: Gusyatin, Mikhail M
; APPLICANT: Rostova, Yulia G
; APPLICANT: Yampolskaya, Tatiana A
; TITLE OF INVENTION: New Mutant Carbamoylphosphate Synthetase and Method for Producing
; FILE REFERENCE: 226350US0
; CURRENT APPLICATION NUMBER: US/10/210,115
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: RU 2001-121697
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 33
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-210-115-33

Query Match 38.0%; Score 57; DB 14; Length 382;
Best Local Similarity 46.7%; Pred. No. 8.6;
Matches 14; Conservative 4; Mismatches 8; Indels 4; Gaps 2;

QY 1 AKE---ASSVDYILG-WFEGGVPEHKKEE 26
Db 159 AKEVTTAEAYSWTQGSWTLTGGLPEAKKD 188

RESULT 15
US-10-369-493-691
; Sequence 691, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 691
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-691

Query Match 38.0%; Score 57; DB 15; Length 382;
Best Local Similarity 46.7%; Pred. No. 8.6;
Matches 14; Conservative 4; Mismatches 8; Indels 4; Gaps 2;

QY 1 AKE---ASSVDYILG-WFEGGVPEHKKEE 26

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:47 ; Search time 43.3719 Seconds
(without alignments)
298.263 Million cell updates/sec

Title: US-10-774-602-14
Perfect score: 204
Sequence: 1 MLSHLYVSKDKENISKEND.....VLDEKEEABETEELK 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	204	100.0	329	5 Q9NFV9	Q9NFV9 plasmodium
2	204	100.0	354	5 Q25995	Q25995 plasmodium
3	204	100.0	354	5 Q81J55	Q81J55 plasmodium
4	204	100.0	361	5 Q95P15	Q95P15 plasmodium
5	204	100.0	379	5 Q25705	Q25705 plasmodium
6	204	100.0	379	5 Q25706	Q25706 plasmodium
7	204	100.0	379	5 Q9U6C4	Q9U6C4 plasmodium
8	204	100.0	380	5 Q26019	Q26019 plasmodium
9	194	95.1	346	5 Q9U0G0	Q9U0G0 plasmodium
10	97	47.5	64	5 Q15762	Q15762 plasmodium
11	94.5	46.3	424	5 Q81J53	Q81J53 plasmodium
12	89	43.6	697	5 Q81J52	Q81J52 plasmodium
13	87	42.6	343	5 Q8SZT9	Q8SZT9 drosophila
14	87	42.6	1147	5 Q9VY73	Q9VY73 drosophila
15	85.5	41.9	406	4 Q8TC47	Q8TC47 homo sapien
16	85.5	41.9	406	4 Q8TC90	Q8TC90 homo sapien

17	84	41.2	405	5 Q81J48	Q81J48 plasmodium
18	82	40.2	1015	5 Q86AG0	Q86AG0 dictyosteli
19	82	40.2	3134	5 Q25994	Q25994 plasmodium
20	82	40.2	3144	5 Q9GTK4	Q9GTK4 plasmodium
21	82	40.2	5890	5 Q81K84	Q81K84 plasmodium
22	81.5	40.0	142	11 Q8C7N3	Q8C7N3 mus musculu
23	80	39.2	232	10 Q9ST20	Q9ST20 brassica ca
24	80	39.2	662	5 Q81AW6	Q81AW6 plasmodium
25	80	39.2	806	5 Q8T2A3	Q8T2A3 dictyosteli
26	80	39.2	874	5 Q81CL6	Q81CL6 plasmodium
27	79.5	39.0	152	10 Q65745	Q65745 cicer ariet
28	79	38.7	204	5 Q81518	Q81518 plasmodium
29	79	38.7	477	13 Q8AYN2	Q8AYN2 gallus gall
30	79	38.7	715	10 Q9LD59	Q9LD59 arabidopsis
31	79	38.7	741	16 Q9JTK3	Q9JTK3 neisseria m
32	79	38.7	1192	4 Q9UL68	Q9UL68 homo sapien
33	79	38.7	1989	5 Q814U7	Q814U7 plasmodium
34	79	38.7	2924	5 Q25733	Q25733 plasmodium
35	78.5	38.5	395	4 Q8NDU3	Q8NDJ3 homo sapien
36	78.5	38.5	695	11 Q8CDJ9	Q8CDJ9 mus musculu
37	78.5	38.5	1331	11 Q9SPQ2	Q9SPQ2 mus musculu
38	78	38.2	331	10 Q8LNT8	Q8LNT8 oryza sativ
39	78	38.2	400	11 Q7TP09	Q7TP09 rattus norv
40	78	38.2	755	11 Q8OW59	Q8OW59 rattus norv
41	78	38.2	772	11 Q8BTT6	Q8BTT6 mus musculu
42	78	38.2	910	12 Q9Q915	Q9Q915 avian adeno
43	78	38.2	1781	4 Q9UKX0	Q9UKX0 homo sapien
44	78	38.2	1890	4 Q9UKW3	Q9UKW3 homo sapien
45	78	38.2	2072	4 Q8WYB5	Q8WYB5 homo sapien

ALIGNMENTS

RESULT 1

ID Q9NFV9 PRELIMINARY; PRT; 329 AA.
AC Q9NFV9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (fragment).
GN MSP3.
OS Plasmodium falciparum (isolate 7G8).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7G8;
RX MEDLINE=20416497; PubMed=10960178;
RA Okenu D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in Plasmodium reichenowi and Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 109:185-188 (2000).
DR EMBL; AJ252287; CAB85901.1; -;
FT NON_TER 1
FT NON_TER 329
SQ SEQUENCE 329 AA; C5B045DB5E21A159 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 329;
Best Local Similarity 100.0%; Pred. No. 5e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLSHLYVSKDKENISKENDVLDKEEAEETEELK 41
Db 201 MLSHLYVSKDKENISKENDVLDKEEAEETEELK 241

RESULT 2

ID Q25995 PRELIMINARY; PRT; 354 AA.
AC Q25995;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

```

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Antigen.
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=NF54;
RC MEDLINE=95198774; PubMed=7891748;
RA McColl D.J., Silva A., Foley M., Kun J.F., Favaloro J.M.,
RA Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;
RT "Molecular variation in a novel polymorphic antigen associated with
RT Plasmodium falciparum merozoites";
RL Mol. Biochem. Parasitol. 68:53-67(1994).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=NF54;
RC MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
RT Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; L28825; AAC09377.1; -.
SQ SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.3e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVDLDEKEEAETEEELEEK 41
|||||
Db 211 MLSHLYVSSKDKENISKENDVDLDEKEEAETEEELEEK 251

RESULT 3
Q8IJ55 PRELIMINARY; PRT; 354 AA.
AC Q8IJ55;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Merozoite surface protein 3.
GN PF10_0345.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=3D7;
RC MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J.J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairhead A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings I.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum";
RL Nature 419:498-511(2002).
DR EMBL; AB014834; AAN35542.1; -.
SQ SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.3e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVDLDEKEEAETEEELEEK 41
|||||
Db 211 MLSHLYVSSKDKENISKENDVDLDEKEEAETEEELEEK 251

RESULT 4
Q95PI5 PRELIMINARY; PRT; 361 AA.
AC Q95PI5;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=FVO;
RC Hiseeda H., Saul A., Long C.A., Miller L.H., Stowers A.W.;
RA "Merozoite Surface Protein 3 and Protection Against Malaria in Aotus
RA Monkeys.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY044180; AAK94780.1; -.
FT NON_TER 1
FT NON_TER 361
SQ SEQUENCE 361 AA; 41163 MW; 6127A3041587BA74 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 361;
Best Local Similarity 100.0%; Pred. No. 5.4e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVDLDEKEEAETEEELEEK 41
|||||
Db 217 MLSHLYVSSKDKENISKENDVDLDEKEEAETEEELEEK 257

RESULT 5
Q25705 PRELIMINARY; PRT; 379 AA.
AC Q25705;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=X1;
RC MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
RT Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; U08851; AAC47831.1; -.
SQ SEQUENCE 379 AA; 43344 MW; DC7AF106887C9AA0 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.6e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVDLDEKEEAETEEELEEK 41
|||||
Db 237 MLSHLYVSSKDKENISKENDVDLDEKEEAETEEELEEK 277

RESULT 6
Q25706 PRELIMINARY; PRT; 379 AA.
AC Q25706;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
```


OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FC27;
 RX MEDLINE=98156743; PubMed=9497029;
 RA McColl D.J., Anders R.F.;
 RT "Conservation of structural motifs and antigenic diversity in the
 Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
 RL Mol. Biochem. Parasitol. 90:21-31(1997).
 DR EMBL; U08852; AAC47832.1; -;
 SQ SEQUENCE 379 AA; 43302 MW; ABF9D54E1ED91A24 CRC64;
 Query Match 100.0%; Score 204; DB 5; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5.6e-12;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLSHLYVSSKDKENISKENDVLDKEEEAEETEELEEK 41
 237 MLSHLYVSSKDKENISKENDVLDKEEEAEETEELEEK 277
 RESULT 7
 Q9U6C4 PRELIMINARY; PRT; 379 AA.
 ID Q9U6C4
 AC Q9U6C4;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Polymorphic antigen.
 GN MSP-3.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FC27;
 RX Li X.R., Yu X.B., Luo S.H., Shan Z.X., Fang J.M., Xu J.;
 RT "Sequence of Plasmodium falciparum secreted polymorphic antigen
 gene.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF188190; AAF04099.1; -;
 SQ SEQUENCE 379 AA; 43316 MW; C152A54E1F9D5F25 CRC64;
 Query Match 100.0%; Score 204; DB 5; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5.6e-12;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLSHLYVSSKDKENISKENDVLDKEEEAEETEELEEK 41
 237 MLSHLYVSSKDKENISKENDVLDKEEEAEETEELEEK 277
 RESULT 8
 Q26019 PRELIMINARY; PRT; 380 AA.
 ID Q26019
 AC Q26019;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Polymorphic antigen precursor.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FC27;
 RX MEDLINE=95198774; PubMed=7891748;
 RA McColl D.J., Silva A., Foley M., Kun J.F., Favaloro J.M.,
 RA Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;
 RT "Molecular variation in a novel polymorphic antigen associated with
 Plasmodium falciparum merozoites.";

RL Mol. Biochem. Parasitol. 68:53-67(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FC27;
 RX MEDLINE=98156743; PubMed=9497029;
 RA McColl D.J., Anders R.F.;
 RT "Conservation of structural motifs and antigenic diversity in the
 Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
 RL Mol. Biochem. Parasitol. 90:21-31(1997).
 DR EMBL; L07944; AAC09378.1; -;
 DR PDB; 1PSM; 07-FEB-95.
 KW SIGNAL.
 FT SIGNAL.
 FT CHAIN 1 25 POTENTIAL.
 SQ SEQUENCE 380 AA; 43290 MW; 0986CA1393094CA2 CRC64;
 Query Match 100.0%; Score 204; DB 5; Length 380;
 Best Local Similarity 100.0%; Pred. No. 5.7e-12;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLSHLYVSSKDKENISKENDVLDKEEEAEETEELEEK 41
 238 MLSHLYVSSKDKENISKENDVLDKEEEAEETEELEEK 278
 RESULT 9
 Q9U0G0 PRELIMINARY; PRT; 346 AA.
 ID Q9U0G0
 AC Q9U0G0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Merozoite surface protein 3 (Pragmat).
 GN MSP3.
 OS Plasmodium reichenowi.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5854;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20416497; PubMed=10960178;
 RA Okenu D.M.N., Thomas A.W., Conway D.J.;
 RT "Allelic lineages of the merozoite surface protein 3 gene in
 Plasmodium reichenowi and Plasmodium falciparum.";
 RL Mol. Biochem. Parasitol. 109:185-188(2000).
 DR EMBL; AJ252286; CAB65754.1; -;
 FT NON_TER 1 1
 FT NON_TER 346 346
 SQ SEQUENCE 346 AA; 39127 MW; A804B96BDFAF010 CRC64;
 Query Match 95.1%; Score 194; DB 5; Length 346;
 Best Local Similarity 95.1%; Pred. No. 4.7e-11;
 Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLSHLYVSSKDKENISKENDVLDKEEEAEETEELEEK 41
 214 MLSHLYVSSKDKENISKENDVLDKEEEAEETEELEEK 254
 RESULT 10
 O15762 PRELIMINARY; PRT; 64 AA.
 ID O15762
 AC O15762;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Merozoite surface protein 3 (Fragment).
 GN MSP-3.
 OS Plasmodium falciparum (isolate NF54).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5843;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NF54;

```

RX MEDLINE=94348085; PubMed=8068948;
RA Oeuvray C., Bouharoun-Tayoun H., Gras-Masse H., Bottius E., Kaidoh T.,
RA Aikawa M., Filgueira M.C., Tartar A., Drullhe P.;
RT "Merozoite surface protein-3: a malaria protein inducing antibodies
RT that promote Plasmodium falciparum killing by cooperation with blood
RT monocytes.";
RL Blood 84:11594-1602(1994) .
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=NF54;
RA Oeuvray C., Bouharoun-Tayoun H., Gras-Masse H., Bottius E., Kaidoh T.,
RA Aikawa M., Filgueira M.-C., Tartar A., Drullhe P.;
RA Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF024624; AAB81610.1; -.
FT NON_TER      1
FT              1
FT NON_TER      64
FT              64
SQ SEQUENCE     64 AA; 7326 MW; 02E0ABF614C15A3B CRC64;

Query Match          47.5%; Score 97; DB 5; Length 64;
Best Local Similarity 95.0%; Pred. No. 0.02;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
   |||||
Db 45 MLSHLYVSSKDKENISKENE 64

RESULT 11
ID Q8IJ53 PRELIMINARY; PRT; 424 AA.
AC Q8IJ53;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF10_0347.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
QC NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Chan J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chen M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Portea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002) .
DR EMBL; AE014834; AAN35544.1; -.
KW Hypothetical protein.
SQ SEQUENCE 424 AA; 48196 MW; 4D2BD2C9A9F3B619 CRC64;

Query Match          46.3%; Score 94.5; DB 5; Length 424;
Best Local Similarity 51.2%; Pred. No. 0.19;
Matches 21; Conservative 10; Mismatches 9; Indels 1; Gaps 1;

QY 1 MLSHLYVSSKDKENISKENDVDLDEKEEEAEETEEELEEK 41
   : : | | | | | | | | | | | | | | | | | | :
Db 285 VLSESVLTSDWKEDIVKENEDVKDEKDDEEBE-EEEKYENE 324

RESULT 12
ID Q8IJ52 PRELIMINARY; PRT; 697 AA.
AC Q8IJ52;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)

```

```

RESULT 14
Q9VVT3 PRELIMINARY; PRT; 1147 AA.
AC Q9VVT3;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
DE 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE CG2025 protein.
GN CG2025.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu P., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fessler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.-H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banon J., An H., Baldwin D., Banon J., Beeson K.Y., Busan D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresner D., Farfan D.,
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,
RA Ibegwam M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;

```

Search completed: September 22, 2004, 18:09:10
Job time : 45.3719 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:00:42 ; Search time 39.876 Seconds
(without alignments)
177.141 Million cell updates/sec

Title: US-10-774-602-11

Perfect score: 124

Sequence: 1 YEKAKYQKQANQAVLKAKESASYD 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Optimal number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	93.5	64	2	AAR54684
2	116	93.5	64	6	ABR41813
3	115	92.7	28	6	ABR41814
4	106	85.5	31	6	ABR41822
5	103	83.1	31	6	ABR41823
6	99	79.8	31	6	ABR41824
7	97	78.2	31	6	ABR41825
8	96	77.4	23	2	AAR54685
9	91	73.4	31	6	ABR41826
10	88	71.0	31	6	ABR41827
11	84	67.7	31	6	ABR41828
12	80	64.5	31	6	ABR41829
13	69	55.6	31	6	ABR41830
14	67	54.0	31	6	ABR41833
15	65	52.4	31	6	ABR41832
16	63	50.8	31	6	ABR41831
17	61	49.2	31	6	ABR41834
18	56	45.2	31	6	ABR41835
19	55	44.4	752	5	ABBS4485
20	53	42.7	463	4	ABBS9725
21	52	41.9	31	6	ABR41836
22	51	41.1	347	3	AAG23630
23	51	41.1	368	3	AAG23629
24	51	41.1	420	3	AAG23628
25	50	40.3	28	2	AAR90177

ALIGNMENTS

RESULT 1

AAR54684
ID AAR54684 standard; protein; 64 AA.

XX AAR54684;

XX 25-MAR-2003 (revised)

DT 15-DEC-1994 (first entry)

XX Plasmodium falciparum merozoite 48kD surface protein fragment I.

XX Merozoite 48kD surface antigen; immunogen; vaccine; malaria;

XX cytophilic antibody.

XX Plasmodium falciparum.

XX WO9409140-A1.

XX 28-APR-1994.

XX 18-OCT-1993; 93WO-FR001024.

XX 19-OCT-1992; 92FR-00012488.

XX (INSP) INST PASTEUR.

XX Druilhe P, Bouharoun-Tayoun H, Oeuvray C;

XX WPI; 1994-151325/18.

XX N-PSDB; AAQ64677.

XX New Plasmodium antigen generating cytophilic - mimicking protection induced by long term exposure to the parasite, useful in vaccination against and diagnosis of malaria, and prodn of antibodies for passive immunisation.

XX Claim 3; Page 36; 52pp; French.

XX This sequence is a preferred fragment of the 48kD merozoite surface antigen of P.falciparum which is useful as an immunogen to generate antibodies (Abs). The Abs can induce a cytotoxic reaction against the erythrocyte stage of the parasite. Specific subfragments of AAR54684 are separately claimed (see AAR54685-R54687). (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 64 AA;

SQ Query Match 93.5%; Score 116; DB 2; Length 64;


```

XX FH Key Location/Qualifiers
XX FT Misc-difference 15..16
XX FT /note= "linker residues"
XX
XX GB2378949-A.
XX
XX 26-FEB-2003.
XX
XX
XX 16-AUG-2001; 2001GB-00020057.
XX
XX (DZIE//) DZIEGIEL M S H.
XX (LUND//) LUNDQUIST R.
XX (NIEL//) NIELSEN L K.
XX
XX Driegiel MSH, Lundquist R, Nielsen LK;
XX WPI; 2003-303108/30.
XX
XX Novel recombinant human antibody specific for merozoite surface protein-3
XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
XX a malarial disease.
XX
XX Example 3; Fig 10; 75pp; English.
XX
XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
XX used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
XX surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
XX acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
XX 15 and 16 replaced by proprietary linker residues. The peptides were
XX examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
XX encoded by clones obtained from a phage display system based on antibody
XX variable region genes from the peripheral blood leucocytes of malaria
XX immune human individuals. RAM2 reacted with the N-terminal, middle and C-
XX terminal part of the MSP-3 antigen, with highest reactivity with the
XX middle part. RAM1 reacted only weakly with the peptides. The invention
XX provides recombinant human antibodies against MSP-3, which are useful for
XX the diagnosis and treatment of malaria. The recombinant human antibodies
XX include the complementarity determining regions from the light and heavy
XX chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
XX
XX Sequence 31 AA;
XX
XX Query Match 85.5%; Score 106; DB 6; Length 31;
XX Best Local Similarity 84.0%; Pred. No. 8.7e-09;
XX Matches 21; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 YERAKNAYQKANOAVLKAKKASSYD 25
XX :|:|||||
XX 1 HERAKNAYQKANOAVLKAKKASSYD 25
XX
XX RESULT 5
XX ABR41823
XX ID ABR41823 standard; peptide; 31 AA.
XX
XX AC ABR41823;
XX
XX 11-AUG-2003 (first entry)
XX
XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
XX
XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
XX immunotherapy; vaccine; epitope.
XX
XX Plasmodium falciparum.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 15..16
XX FT /note= "linker residues"
XX
XX
XX GB2378949-A.
XX
XX 26-FEB-2003.
XX
XX
XX 16-AUG-2001; 2001GB-00020057.
XX
XX (DZIE//) DZIEGIEL M S H.
XX (LUND//) LUNDQUIST R.
XX (NIEL//) NIELSEN L K.
XX
XX Driegiel MSH, Lundquist R, Nielsen LK;
XX WPI; 2003-303108/30.
XX
XX Novel recombinant human antibody specific for merozoite surface protein-3
XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
XX a malarial disease.
XX
XX Example 3; Fig 10; 75pp; English.
XX
XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
XX used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
XX surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
XX acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
XX 15 and 16 replaced by proprietary linker residues. The peptides were
XX examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
XX encoded by clones obtained from a phage display system based on antibody
XX variable region genes from the peripheral blood leucocytes of malaria
XX immune human individuals. RAM2 reacted with the N-terminal, middle and C-
XX terminal part of the MSP-3 antigen, with highest reactivity with the
XX middle part. RAM1 reacted only weakly with the peptides. The invention
XX provides recombinant human antibodies against MSP-3, which are useful for
XX the diagnosis and treatment of malaria. The recombinant human antibodies
XX include the complementarity determining regions from the light and heavy
XX chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
XX
XX Sequence 31 AA;
XX
XX Query Match 85.5%; Score 106; DB 6; Length 31;
XX Best Local Similarity 84.0%; Pred. No. 8.7e-09;
XX Matches 21; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 YERAKNAYQKANOAVLKAKKASSYD 25
XX :|:|||||
XX 1 HERAKNAYQKANOAVLKAKKASSYD 25
XX
XX RESULT 6
XX ABR41824
XX ID ABR41824 standard; peptide; 31 AA.
XX
XX AC ABR41824;
XX
XX 11-AUG-2003 (first entry)
XX
XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
XX
XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
XX immunotherapy; vaccine; epitope.
XX
XX Plasmodium falciparum.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 15..16
XX FT /note= "linker residues"
XX
XX
XX GB2378949-A.
XX
XX 26-FEB-2003.
XX
XX
XX 16-AUG-2001; 2001GB-00020057.
XX
XX (DZIE//) DZIEGIEL M S H.
XX (LUND//) LUNDQUIST R.
XX (NIEL//) NIELSEN L K.
XX
XX Driegiel MSH, Lundquist R, Nielsen LK;
XX WPI; 2003-303108/30.
XX
XX Novel recombinant human antibody specific for merozoite surface protein-3
XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
XX a malarial disease.
XX
XX Example 3; Fig 10; 75pp; English.
XX
XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
XX used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
XX surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
XX acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
XX 15 and 16 replaced by proprietary linker residues. The peptides were
XX examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
XX encoded by clones obtained from a phage display system based on antibody
XX variable region genes from the peripheral blood leucocytes of malaria
XX immune human individuals. RAM2 reacted with the N-terminal, middle and C-
XX terminal part of the MSP-3 antigen, with highest reactivity with the
XX middle part. RAM1 reacted only weakly with the peptides. The invention
XX provides recombinant human antibodies against MSP-3, which are useful for
XX the diagnosis and treatment of malaria. The recombinant human antibodies
XX include the complementarity determining regions from the light and heavy
XX chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
XX
XX Sequence 31 AA;
XX
XX Query Match 83.1%; Score 103; DB 6; Length 31;
XX Best Local Similarity 87.5%; Pred. No. 2.5e-08;
XX Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX 2 EKAKNAYQKANOAVLKAKKASSYD 25
XX :|:|||||
XX 1 ERAKNAYQKANOAVXXKAKKASSYD 24
XX
XX RESULT 6
XX ABR41824
XX ID ABR41824 standard; peptide; 31 AA.
XX
XX AC ABR41824;
XX
XX 11-AUG-2003 (first entry)
XX
XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
XX
XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
XX immunotherapy; vaccine; epitope.
XX
XX Plasmodium falciparum.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 15..16
XX FT /note= "linker residues"
XX
XX
XX GB2378949-A.
XX
XX 26-FEB-2003.
XX
XX
XX 16-AUG-2001; 2001GB-00020057.

```

```

XX 16-AUG-2001; 2001GB-00020057.
XX PR
XX (DZIE/) DZIEGIEL M S H.
XX PA
XX (LUND/) LUNDQUIST R.
XX PA
XX (NIEL/) NIELSEN L K.
XX PI
XX Dziegiel MSH, Lundquist R, Nielsen LK;
XX WPI; 2003-303108/30.
XX PT
XX Novel recombinant human antibody specific for merozoite surface protein-3
XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
XX a malarial disease.
XX PS
XX Example 3; Fig 10; 75pp; English.
XX CC
XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
XX used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
XX surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
XX acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
XX 15 and 16 replaced by proprietary linker residues. The peptides were
XX examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
XX encoded by clones obtained from a phage display system based on antibody
XX variable region genes from the peripheral blood leucocytes of malaria
XX immune human individuals. RAM2 reacted with the N-terminal, middle and C-
XX terminal part of the MSP-3 antigen, with highest reactivity with the
XX middle part. RAM1 reacted only weakly with the peptides. The invention
XX provides recombinant human antibodies against MSP-3, which are useful for
XX the diagnosis and treatment of malaria. The recombinant human antibodies
XX include the complementarity determining regions from the light and heavy
XX chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
XX SQ
XX Sequence 31 AA;
XX Query Match 79.8%; Score 99; DB 6; Length 31;
XX Best Local Similarity 87.0%; Pred. No. 9.9e-08;
XX Matches 20; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 3 KAKNAYQKANQAVLKAKKEASSYD 25
XX :|||||
XX 1 RAKNAYQKANQAVLKXKEASSYD 23
XX Db
XX
XX RESULT 7
XX ABR41825
XX ABR41825 standard; peptide; 31 AA.
XX ABR41825;
XX DT 11-AUG-2003 (first entry)
XX DE Plasmodium falciparum merozoite 48kD surface protein fragment II.
XX KW Merozoite 48kD surface antigen; immunogen; vaccine; malaria;
XX OS Merozoite surface protein; MSP-3; malaria; antimalarial; protozoaside;
XX KW immunotherapy; vaccine; epitope.
XX OS Plasmodium falciparum.
XX XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 15..16
XX FT /note= "linker residues"
XX XX
XX PN GB2378949-A.
XX XX
XX PD 26-FEB-2003.
XX XX
XX PF 16-AUG-2001; 2001GB-00020057.
XX XX
XX PR 16-AUG-2001; 2001GB-00020057.
XX XX
XX (DZIE/) DZIEGIEL M S H.
XX PA
XX (LUND/) LUNDQUIST R.
XX PA

```

```

PA (NIEL/) NIELSEN L K.
XX Dziegiel MSH, Lundquist R, Nielsen LK;
XX WPI; 2003-303108/30.
XX PT
XX Novel recombinant human antibody specific for merozoite surface protein-3
XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
XX a malarial disease.
XX PS
XX Example 3; Fig 10; 75pp; English.
XX CC
XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
XX used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
XX surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
XX acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
XX 15 and 16 replaced by proprietary linker residues. The peptides were
XX examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
XX encoded by clones obtained from a phage display system based on antibody
XX variable region genes from the peripheral blood leucocytes of malaria
XX immune human individuals. RAM2 reacted with the N-terminal, middle and C-
XX terminal part of the MSP-3 antigen, with highest reactivity with the
XX middle part. RAM1 reacted only weakly with the peptides. The invention
XX provides recombinant human antibodies against MSP-3, which are useful for
XX the diagnosis and treatment of malaria. The recombinant human antibodies
XX include the complementarity determining regions from the light and heavy
XX chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
XX SQ
XX Sequence 31 AA;
XX Query Match 78.2%; Score 97; DB 6; Length 31;
XX Best Local Similarity 90.9%; Pred. No. 2e-07;
XX Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 4 AKNAYQKANQAVLKAKKEASSYD 25
XX :|||||
XX 1 AKNAYQKANQAVLKXKEASSYD 22
XX Db
XX
XX RESULT 8
XX AAR54685
XX ID AAR54685 standard; protein; 23 AA.
XX AC AAR54685;
XX XX
XX DT 25-MAR-2003 (revised)
XX DT 15-DEC-1994 (first entry)
XX DE Plasmodium falciparum merozoite 48kD surface protein fragment II.
XX KW Merozoite 48kD surface antigen; immunogen; vaccine; malaria;
XX KW Cytophilic antibody.
XX OS Plasmodium falciparum.
XX OS Plasmodium falciparum.
XX PN WO9409140-A1.
XX XX
XX PD 28-APR-1994.
XX XX
XX PF 18-OCT-1993; 93WO-FR001024.
XX XX
XX PR 19-OCT-1992; 92FR-00012488.
XX XX
XX PA (INSP ) INST PASTEUR.
XX PI Druilhe P, Bouharoun-Tayoun H, Oeuvray C;
XX WPI; 1994-151325/18.
XX DR
XX New Plasmodium antigen generating cytophilic - mimicking protection
XX induced by long term exposure to the parasite, useful in vaccination
XX against and diagnosis of malaria, and prodn of antibodies for passive
XX immunisation.
XX PT

```


XX Claim 4; Page 37; 52pp; French.

XX This sequence is a preferred fragment of the 48kD merozoite surface

CC antigen of P. falciparum which is useful as an immunogen to generate

CC antibodies (Abs). The Abs can induce a cytotoxic reaction against the

CC erythrocyte stage of the parasite. (Updated on 25-MAR-2003 to correct PN

CC field.)

XX Sequence 23 AA;

XX Query Match 77.4%; Score 96; DB 2; Length 23;

XX Best Local Similarity 84.0%; Pred. No. 2e-07; Indels 2; Gaps 1;

XX Matches 21; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY 1 YEKAKNAYQKQAVLKAKKASSYD 25

DB 1 HERAKNAYQKQAVLKAKKASSYD 23

RESULT 9

ABR41826

ABR41826 standard; peptide; 31 AA.

AC ABR41826;

XX 11-AUG-2003 (first entry)

XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;

KW immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 15.16

FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE//) DZIEGIEL M S H.

XX (LUND//) LUNDQUIST R.

XX (NIEL//) NIELSEN L K.

PI Driegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3

XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of

XX a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)

XX used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite

XX surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino

XX acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues

XX 15 and 16 replaced by proprietary linker residues. The peptides were

XX examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are

XX encoded by clones obtained from a phage display system based on antibody

XX variable region genes from the peripheral blood leucocytes of malaria

XX immune human individuals. RAM2 reacted with the N-terminal, middle and C-

XX terminal part of the MSP-3 antigen, with highest reactivity with the

XX middle part. RAM1 reacted only weakly with the peptides. The invention

XX provides recombinant human antibodies against MSP-3, which are useful for

XX the diagnosis and treatment of malaria. The recombinant human antibodies

XX include the complementarity determining regions from the light and heavy

XX chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

CC the diagnosis and treatment of malaria. The recombinant human antibodies

CC include the complementarity determining regions from the light and heavy

CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

XX Query Match 73.4%; Score 91; DB 6; Length 31;

XX Best Local Similarity 90.5%; Pred. No. 1.6e-06;

XX Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 KNAYQKQAVLKAKKASSYD 25

DB 1 KNAYQKQAVLKAKKASSYD 21

RESULT 10

ABR41827

ABR41827 standard; peptide; 31 AA.

AC ABR41827;

XX 11-AUG-2003 (first entry)

XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;

KW immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 15.16

FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE//) DZIEGIEL M S H.

XX (LUND//) LUNDQUIST R.

XX (NIEL//) NIELSEN L K.

PI Driegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3

XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of

XX a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)

XX used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite

XX surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino

XX acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues

XX 15 and 16 replaced by proprietary linker residues. The peptides were

XX examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are

XX encoded by clones obtained from a phage display system based on antibody

XX variable region genes from the peripheral blood leucocytes of malaria

XX immune human individuals. RAM2 reacted with the N-terminal, middle and C-

XX terminal part of the MSP-3 antigen, with highest reactivity with the

XX middle part. RAM1 reacted only weakly with the peptides. The invention

XX provides recombinant human antibodies against MSP-3, which are useful for

XX the diagnosis and treatment of malaria. The recombinant human antibodies

XX include the complementarity determining regions from the light and heavy

XX chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 71.0%; Score 88; DB 6; Length 31;
 Best Local Similarity 90.0%; Pred. No. 4.5e-06;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 NAYQKANOAVLKAKEASSYD 25
 DB 1 NAYQKANOAVLKAKEASSYD 20

RESULT 11
 ABR41828
 ID ABR41828 standard; peptide; 31 AA.
 AC ABR41828;
 XX
 XX DT 11-AUG-2003 (first entry)
 XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX Plasmodium falciparum.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 15..16
 FT /note= "linker residues"
 FT
 XX GB2378949-A.
 PN
 XX 26-FEB-2003.
 PD
 XX 16-AUG-2001; 2001GB-00020057.
 PF
 XX 16-AUG-2001; 2001GB-00020057.
 PR
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX
 XX Dziegiel MSH, Lundquist R, Nielsen LK;
 PI WPI; 2003-303108/30.
 DR
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

Example 3; Fig 10; 75pp; English.
 The present sequence is one of a set of 34 peptides (see ABR41822-55)
 used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 15 and 16 replaced by proprietary linker residues. The peptides were
 examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 encoded by clones obtained from a phage display system based on antibody
 variable region genes from the peripheral blood leucocytes of malaria
 immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 terminal part of the MSP-3 antigen, with highest reactivity with the
 middle part. RAM1 reacted only weakly with the peptides. The invention
 provides recombinant human antibodies against MSP-3, which are useful for
 the diagnosis and treatment of malaria. The recombinant human antibodies
 include the complementarity determining regions from the light and heavy
 chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

Sequence 31 AA;
 Query Match 67.7%; Score 84; DB 6; Length 31;
 Best Local Similarity 89.5%; Pred. No. 1.8e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 AYQKANOAVLKAKEASSYD 25
 DB 1 AYQKANOAVLKAKEASSYD 19

RESULT 12
 ABR41829
 ID ABR41829 standard; peptide; 31 AA.
 AC ABR41829;
 XX
 XX DT 11-AUG-2003 (first entry)
 XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX Plasmodium falciparum.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 15..16
 FT /note= "linker residues"
 FT
 XX GB2378949-A.
 PN
 XX 26-FEB-2003.
 PD
 XX 16-AUG-2001; 2001GB-00020057.
 PF
 XX 16-AUG-2001; 2001GB-00020057.
 PR
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX
 XX Dziegiel MSH, Lundquist R, Nielsen LK;
 PI WPI; 2003-303108/30.
 DR
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

Example 3; Fig 10; 75pp; English.
 The present sequence is one of a set of 34 peptides (see ABR41822-55)
 used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 15 and 16 replaced by proprietary linker residues. The peptides were
 examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 encoded by clones obtained from a phage display system based on antibody
 variable region genes from the peripheral blood leucocytes of malaria
 immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 terminal part of the MSP-3 antigen, with highest reactivity with the
 middle part. RAM1 reacted only weakly with the peptides. The invention
 provides recombinant human antibodies against MSP-3, which are useful for
 the diagnosis and treatment of malaria. The recombinant human antibodies
 include the complementarity determining regions from the light and heavy
 chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

Sequence 31 AA;
 Query Match 64.5%; Score 80; DB 6; Length 31;
 Best Local Similarity 88.9%; Pred. No. 7.2e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 YOKANOAVLKAKEASSYD 25
 DB 1 YOKANOAVLKAKEASSYD 18

RESULT 13
 ABR41830
 ID ABR41830 standard; peptide; 31 AA.
 AC ABR41830;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX
 KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX
 OS Plasmodium falciparum.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 15. .16
 FT /note= "linker residues"
 XX
 PN GB2378949-A.
 XX
 PD 26-FEB-2003.
 XX
 PF 16-AUG-2001; 2001GB-00020057.
 XX
 PR 16-AUG-2001; 2001GB-00020057.
 XX
 PA (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX
 PI Driegiel MSH, Lundquist R, Nielsen LK;
 XX
 DR WPI; 2003-303108/30.
 XX
 PT Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX
 PS Example 3; Fig 10; 75pp; English.
 XX
 CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX
 SQ Sequence 31 AA;
 Query Match 55.6%; Score 69; DB 6; Length 31;
 Best Local Similarity 88.2%; Pred. No. 0.0033;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 9 QKANQAVLKAKESASYD 25
 DB 1 QKANQAVLKAKESASYD 17
 XX
 AC ABR41833
 ID ABR41833 standard; peptide; 31 AA.
 XX
 AC ABR41833;
 XX

XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX
 KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX
 OS Plasmodium falciparum.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 15. .16
 FT /note= "linker residues"
 XX
 PN GB2378949-A.
 XX
 PD 26-FEB-2003.
 XX
 PF 16-AUG-2001; 2001GB-00020057.
 XX
 PR 16-AUG-2001; 2001GB-00020057.
 XX
 PA (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX
 PI Driegiel MSH, Lundquist R, Nielsen LK;
 XX
 DR WPI; 2003-303108/30.
 XX
 PT Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX
 PS Example 3; Fig 10; 75pp; English.
 XX
 CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX
 SQ Sequence 31 AA;
 Query Match 54.0%; Score 67; DB 6; Length 31;
 Best Local Similarity 100.0%; Pred. No. 0.0065;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 NQAVLKAKESASYD 25
 DB 1 NQAVLKAKESASYD 14
 XX
 AC ABR41832
 ID ABR41832 standard; peptide; 31 AA.
 XX
 AC ABR41832;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX

KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX
 OS Plasmodium falciparum.
 XX
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 15..16
 FT /note= "linker residues"
 XX
 XX
 PN GB2378949-A.
 XX
 XX
 PD 26-FEB-2003.
 XX
 XX
 PF 16-AUG-2001; 2001GB-00020057.
 XX
 PF 16-AUG-2001; 2001GB-00020057.
 PR
 XX
 PA (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 PI Dzigiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX
 DR
 XX
 PT Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX
 XX
 PS Example 3; Fig 10; 75pp; English.
 XX
 XX
 CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IqG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

Sequence 31 AA;
 Query Match 52.4%; Score 65; DB 6; Length 31;
 Best Local Similarity 100.0%; Pred. No. 0.013;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ANQAVLKAKEASSY 24
 |||||
 Db 1 ANQAVLKAKEASSY 14

Search completed: September 22, 2004, 18:06:19
 Job time : 41.876 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:00:42 ; Search time 43.0661 Seconds
(without alignments)
177.141 Million cell updates/sec

Title: US-10-774-602-12
Perfect score: 150
Sequence: 1 AKRASSYDYLWFGGVEGHKKEEN 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	100.0	27	2	AAR54686 Plasmodium
2	150	100.0	27	6	ABR41815 Merozoite
3	150	100.0	64	2	AAR54684 Plasmodium
4	150	100.0	64	6	ABR41813 Merozoite
5	138	92.0	31	6	ABR41835 Peptide u
6	137	91.3	31	6	ABR41838 Peptide u
7	136	90.7	31	6	ABR41839 Peptide u
8	134	89.3	31	6	ABR41834 Peptide u
9	132	88.0	31	6	ABR41840 Peptide u
10	131	87.3	31	6	ABR41837 Peptide u
11	130	86.7	31	6	ABR41836 Peptide u
12	127	84.7	31	6	ABR41841 Peptide u
13	126	84.0	31	6	ABR41833 Peptide u
14	124	82.7	31	6	ABR41842 Peptide u
15	119	79.3	31	6	ABR41832 Peptide u
16	118	78.7	31	6	ABR41843 Peptide u
17	114	76.0	31	6	ABR41831 Peptide u
18	113	75.3	31	6	ABR41844 Peptide u
19	112	74.7	31	6	ABR41830 Peptide u
20	109	72.7	31	6	ABR41845 Peptide u
21	108	72.0	31	6	ABR41829 Peptide u
22	103	68.7	31	6	ABR41828 Peptide u
23	102	68.0	31	6	ABR41846 Peptide u
24	99	66.0	31	6	ABR41847 Peptide u
25	94	62.7	31	6	ABR41827 Peptide u

26	92	61.3	31	6	ABR41848	AbR41848 Peptide u
27	88	58.7	31	6	ABR41849	AbR41849 Peptide u
28	88	58.7	31	6	ABR41826	AbR41826 Peptide u
29	85	56.7	31	6	ABR41850	AbR41850 Peptide u
30	85	56.7	31	6	ABR41825	AbR41825 Peptide u
31	84	56.0	31	6	ABR41851	AbR41851 Peptide u
32	84	56.0	31	6	ABR41824	AbR41824 Peptide u
33	82	54.7	31	6	ABR41823	AbR41823 Peptide u
34	79	52.7	31	6	ABR41852	AbR41852 Peptide u
35	76	50.7	31	6	ABR41822	AbR41822 Peptide u
36	74	49.3	31	6	ABR41853	AbR41853 Peptide u
37	68	45.3	31	6	ABR41854	AbR41854 Peptide u
38	65	43.3	11	6	ABR41856	AbR41856 Merozoite
39	62	41.3	31	6	ABR41855	AbR41855 Peptide u
40	58	38.7	388	6	ABMG7504	Abmg7504 Phototab
41	58	38.7	596	6	ABU23235	Abu23235 Protein e
42	58	38.7	599	6	ABU21874	Abu21874 Protein e
43	58	38.7	600	6	ABU20096	Abu20096 Protein e
44	58	38.7	600	6	ABU22081	Abu22081 Protein e
45	57	38.0	382	6	ABU15372	Abu15372 Protein e

ALIGNMENTS

RESULT 1
AAR54686
ID AAR54686 standard; protein; 27 AA.

AC AAR54686;
XX
XX
DT 25-MAR-2003 (revised)
DT 15-DEC-1994 (first entry)
XX
DE Plasmodium falciparum merozoite 48kD surface protein fragment III.
XX
KW Merozoite 48kD surface antigen; immunogen; vaccine; malaria;
KW cytophilic antibody.
XX
OS Plasmodium falciparum.
XX
PN WO9409140-A1.
XX
PD 28-APR-1994.
XX
PF 18-OCT-1993; 93WO-FR001024.
XX
PR 19-OCT-1992; 92FR-00012488.
(INSP) INST PASTEUR.
Druilhe P, Bouharoun-Tayoun H, Oeuvery C;
WPI; 1994-151325/18.
N-PSDB; AAQ64679.
New Plasmodium antigen generating cytophilic - mimicking protection
induced by long term exposure to the parasite, useful in vaccination
against and diagnosis of malaria, and prodn of antibodies for passive
immunisation.
Claim 4; Page 37; 52pp; French.

CC This sequence is a preferred fragment of the 48kD merozoite surface
antigen of P.falciparum which is useful as an immunogen to generate
antibodies (Abs). The Abs can induce a cytotoxic reaction against the
erythrocyte stage of the parasite. (Updated on 25-MAR-2003 to correct PN
field.)
XX
SQ Sequence 27 AA;
Query Match 100.0%; Score 150; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.7e-15;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPHEHKEEN 27
 DB 1 AKEASSYDYLGWFGGVPHEHKEEN 27

RESULT 2
 ABR41815

ID ABR41815 standard; protein; 27 AA.
 AC ABR41815;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Merozoite surface antigen MSP-3 peptide (aa211-237).
 XX
 KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 XX immunotherapy; vaccine.
 OS Plasmodium falciparum.
 XX

Key Location/Qualifiers
 FH 10..20
 FT /note= "epitope specifically referred to in Claim 3"
 XX

PN GB2378949-A.
 XX
 PD 26-FEB-2003.
 XX
 PF 16-AUG-2001; 2001GB-00020057.
 XX
 PR 16-AUG-2001; 2001GB-00020057.
 XX

(DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX
 PI Dzigiel MSH, Lundquist R, Nielsen LK;
 XX
 DR WPI; 2003-303108/30.
 XX
 PT Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX

PS Disclosure; Fig 3; 75pp; English.
 XX

CC The present sequence is that of amino acids 211-237 of the merozoite
 CC surface protein 3 (MSP-3) of Plasmodium falciparum clone D10. The
 CC invention relates to the construction of recombinant human antibodies
 CC against MSP-3, in particular MSP-3(194-257) antigen (see ABR41813). The
 CC antibodies (see also ABR41816-21) are able to passively induce naturally
 CC occurring non-sterile malaria immunity (termed premunition) and exert
 CC their effect via the antibody-dependent cellular inhibition of growth
 CC (ADCI) mechanism. They can be used in the diagnosis and treatment of
 CC malaria. The invention also provides a method of screening a phage
 CC library with target molecules, such as MSP-3 antigen, bound to beads,
 CC preferably magnetic microspheres
 XX
 SQ Sequence 27 AA;

Query Match 100.0%; Score 150; DB 6; Length 27;
 Best Local Similarity 100.0%; Pred. No. 5.7e-15;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPHEHKEEN 27
 DB 1 AKEASSYDYLGWFGGVPHEHKEEN 27

RESULT 3
 AAR54684

ID AAR54684 standard; protein; 64 AA.
 XX
 AC AAR54684;
 XX
 DT 25-MAR-2003 (revised)
 DT 15-DEC-1994 (first entry)
 XX
 DE Plasmodium falciparum merozoite 48kD surface protein fragment I.
 XX
 KW Merozoite 48kD surface antigen; immunogen; vaccine; malaria;
 KW cytophilic antibody.
 OS Plasmodium falciparum.
 XX
 PN WO9409140-A1.
 XX
 PD 28-APR-1994.
 XX
 PF 18-OCT-1993; 93WO-FR001024.
 XX
 PR 19-OCT-1992; 92FR-00012488.
 XX

(INSP) INST PASTEUR.
 PA
 PI Druilhe P, Bouharoun-Tayoun H, Oeuvray C;
 XX
 DR WPI; 1994-151325/18.
 DR N-PSDB; AAQ64677.
 XX

New Plasmodium antigen generating cytophilic - mimicking protection
 PT induced by long term exposure to the parasite, useful in vaccination
 PT against and diagnosis of malaria, and prodn of antibodies for passive
 PT immunisation.
 XX

PS Claim 3; Page 36; 52pp; French.
 XX

CC This sequence is a preferred fragment of the 48kD merozoite surface
 CC antigen of P.falciparum which is useful as an immunogen to generate
 CC antibodies (Abs). The Abs can induce a cytotoxic reaction against the
 CC erythrocyte stage of the parasite. Specific subfragments of AAR54684 are
 CC separately claimed (see AAR54685-R54687). (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX

SQ Sequence 64 AA;

Query Match 100.0%; Score 150; DB 2; Length 64;
 Best Local Similarity 100.0%; Pred. No. 1.4e-14;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPHEHKEEN 27
 DB 18 AKEASSYDYLGWFGGVPHEHKEEN 44

RESULT 4
 ABR41813

ID ABR41813 standard; protein; 64 AA.
 XX
 AC ABR41813;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Merozoite surface antigen MSP-3 peptide (aa194-257).
 XX
 KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine.
 XX
 OS Plasmodium falciparum.
 XX

Key Location/Qualifiers
 FH 27..37
 FT /note= "epitope specifically referred to in Claim 3"
 XX

PN GB2378949-A.
 XX 26-FEB-2003.
 XX 16-AUG-2001; 2001GB-00020057.
 XX 16-AUG-2001; 2001GB-00020057.
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX Driegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX Example 3; Fig 3; 75pp; English.
 XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX Sequence 64 AA;
 SQ Query Match 100.0%; Score 150; DB 6; Length 64;
 Best Local Similarity 100.0%; Pred. No. 1.4e-14;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AKEASSYDYILGWFGGVPPEHKKEEN 27
 DB 18 AKEASSYDYILGWFGGVPPEHKKEEN 44
 RESULT 5
 ABR41835
 ID ABR41835 standard; peptide; 31 AA.
 XX ABR41835;
 DT 11-AUG-2003 (first entry)
 XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 DE Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 KW Plasmodium falciparum.
 OS Key Location/Qualifiers
 XX Key Location/Qualifiers
 FH Misc-difference 15..16
 FT /note= "linker residues"
 FT
 XX GB2378949-A.
 XX 26-FEB-2003.
 XX 16-AUG-2001; 2001GB-00020057.
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX Driegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX Example 3; Fig 3; 75pp; English.
 XX The present sequence is that of amino acids 194-257 of the merozoite
 CC surface protein 3 (MSP-3) of Plasmodium falciparum clone D10. The
 CC invention relates to the construction of recombinant human antibodies
 CC against MSP-3, in particular MSP-3(194-257) antigen. The antibodies (see
 CC also ABR41816-21) are able to passively induce naturally occurring non-
 CC sterile malaria immunity (termed premunition) and exert their effect via
 CC the antibody-dependent cellular inhibition of growth (ADCI) mechanism.
 CC They can be used in the diagnosis and treatment of malaria. The invention
 CC also provides a method of screening a phage library with target
 CC molecules, such as MSP-3 antigen, bound to beads, preferably magnetic
 CC microspheres
 XX Sequence 64 AA;
 SQ Query Match 100.0%; Score 150; DB 6; Length 64;
 Best Local Similarity 100.0%; Pred. No. 1.4e-14;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AKEASSYDYILGWFGGVPPEHKKEEN 27
 DB 18 AKEASSYDYILGWFGGVPPEHKKEEN 44
 RESULT 5
 ABR41835
 ID ABR41835 standard; peptide; 31 AA.
 XX ABR41835;
 DT 11-AUG-2003 (first entry)
 XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 DE Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 KW Plasmodium falciparum.
 OS Key Location/Qualifiers
 XX Key Location/Qualifiers
 FH Misc-difference 15..16
 FT /note= "linker residues"
 FT
 XX GB2378949-A.
 XX 26-FEB-2003.
 XX 16-AUG-2001; 2001GB-00020057.
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX Driegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX Example 3; Fig 3; 75pp; English.
 XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX Sequence 31 AA;
 SQ Query Match 92.0%; Score 138; DB 6; Length 31;
 Best Local Similarity 92.6%; Pred. No. 4e-13;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AKEASSYDYILGWFGGVPPEHKKEEN 27
 DB 5 AKEASSYDYIXXWFGGVPPEHKKEEN 31
 RESULT 6
 ABR41838
 ID ABR41838 standard; peptide; 31 AA.
 XX ABR41838;
 DT 11-AUG-2003 (first entry)
 XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 DE Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 KW Plasmodium falciparum.
 OS Key Location/Qualifiers
 XX Key Location/Qualifiers
 FH Misc-difference 15..16
 FT /note= "linker residues"
 FT
 XX GB2378949-A.
 XX 26-FEB-2003.
 XX 16-AUG-2001; 2001GB-00020057.
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX Driegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX Example 3; Fig 3; 75pp; English.
 XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX Sequence 31 AA;
 SQ Query Match 92.0%; Score 138; DB 6; Length 31;
 Best Local Similarity 92.6%; Pred. No. 4e-13;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AKEASSYDYILGWFGGVPPEHKKEEN 27
 DB 5 AKEASSYDYIXXWFGGVPPEHKKEEN 31

DR WPI; 2003-303108/30.
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX
 XX Example 3; Fig 10; 75pp; English.
 PS
 PS The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX
 XX Sequence 31 AA;

Query Match 91.3%; Score 137; DB 6; Length 31;
 Best Local Similarity 92.6%; Pred. No. 5.6e-13;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPHEKKEN 27
 DB 2 AKEASSYDYLGWXXGGVPEHKKEN 28

RESULT 7
 ABR41839
 ID ABR41839 standard; peptide; 31 AA.
 XX
 AC ABR41839;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX
 KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX
 OS Plasmodium falciparum.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 15..16 /note= "linker residues"
 FT
 FT
 FT
 PN GB2378949-A.
 XX
 PD 26-FEB-2003.
 XX
 PF 16-AUG-2001; 2001GB-00020057.
 XX
 PR 16-AUG-2001; 2001GB-00020057.
 XX
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX
 PI Dziegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

XX
 PS Example 3; Fig 10; 75pp; English.
 XX
 CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX
 XX Sequence 31 AA;

Query Match 90.7%; Score 136; DB 6; Length 31;
 Best Local Similarity 92.6%; Pred. No. 7.9e-13;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPHEKKEN 27
 DB 1 AKEASSYDYLGWXXGGVPEHKKEN 27

RESULT 8
 ABR41834
 ID ABR41834 standard; peptide; 31 AA.
 XX
 AC ABR41834;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX
 KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX
 OS Plasmodium falciparum.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 15..16 /note= "linker residues"
 FT
 FT
 FT
 PN GB2378949-A.
 XX
 PD 26-FEB-2003.
 XX
 PF 16-AUG-2001; 2001GB-00020057.
 XX
 PR 16-AUG-2001; 2001GB-00020057.
 XX
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX
 PI Dziegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX
 PS Example 3; Fig 10; 75pp; English.
 XX
 CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite

CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX
 SQ Sequence 31 AA;

Query Match 89.3%; Score 134; DB 6; Length 31;
 Best Local Similarity 92.3%; Pred. No. 1.6e-12;
 Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 DB
 1 AKEASSYDYILGWFGGVPPEHKKEE 26
 6 AKEASSYDYXXGWFGGVPPEHKKEE 31

RESULT 9
 ABR41840
 ID ABR41840 standard; peptide; 31 AA.
 XX
 AC ABR41840;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX
 KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX
 OS Plasmodium falciparum.

Key Location/Qualifiers
 FT Misc-difference 15..16 /note= "linker residues"
 FT

GB2378949-A.
 26-FEB-2003.

16-AUG-2001; 2001GB-00020057.
 16-AUG-2001; 2001GB-00020057.

(DZIE/) DZIEGIEL M S H.
 (LUND/) LUNDQUIST R.
 (NIEL/) NIELSEN L K.

Dziegiel MSH, Lundquist R, Nielsen LK;
 WPI; 2003-303108/30.

Novel recombinant human antibody specific for merozoite surface protein-3
 antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 a malarial disease.

Example 3; Fig 10; 75pp; English.

The present sequence is one of a set of 34 peptides (see ABR41822-55)
 used in a pepSCAN epitope mapping of Plasmodium falciparum merozoite
 surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody

CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX
 SQ Sequence 31 AA;

Query Match 88.0%; Score 132; DB 6; Length 31;
 Best Local Similarity 92.3%; Pred. No. 3.1e-12;
 Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEASSYDYILGWFGGVPPEHKKEEN 27
 DB 1 KEASSYDYILGWFGGVPPEHKKEEN 26

RESULT 10
 ABR41837
 ID ABR41837 standard; peptide; 31 AA.
 XX
 AC ABR41837;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.

Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 immunotherapy; vaccine; epitope.

Plasmodium falciparum.

Key Location/Qualifiers
 FT Misc-difference 15..16 /note= "linker residues"
 FT

GB2378949-A.
 26-FEB-2003.

16-AUG-2001; 2001GB-00020057.
 16-AUG-2001; 2001GB-00020057.

(DZIE/) DZIEGIEL M S H.
 (LUND/) LUNDQUIST R.
 (NIEL/) NIELSEN L K.

Dziegiel MSH, Lundquist R, Nielsen LK;
 WPI; 2003-303108/30.

Novel recombinant human antibody specific for merozoite surface protein-3
 antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 a malarial disease.

Example 3; Fig 10; 75pp; English.

The present sequence is one of a set of 34 peptides (see ABR41822-55)
 used in a pepSCAN epitope mapping of Plasmodium falciparum merozoite
 surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for

CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX

SQ Sequence 31 AA;

Query Match 87.3%; Score 131; DB 6; Length 31;
 Best Local Similarity 92.6%; Pred. No. 4.3e-12;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPPEHKKEEN 27
 |||||
 Db 3 AKEASSYDYLGXFGGVPPEHKKEEN 29
 |||||

RESULT 11
 ABR41836
 ID ABR41836 standard; peptide; 31 AA.

XX ABR41836;

DT 11-AUG-2003 (first entry)

XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 15..16
 FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

XX (LUND/) LUNDQUIST R.

XX (NIEL/) NIELSEN L K.

XX Dziegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 86.7%; Score 130; DB 6; Length 31;
 Best Local Similarity 92.6%; Pred. No. 6.1e-12;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPPEHKKEEN 27
 |||||
 Db 4 AKEASSYDYLXXEFGGVPPEHKKEEN 30
 |||||

RESULT 12
 ABR41841

ID ABR41841 standard; peptide; 31 AA.

XX ABR41841;

DT 11-AUG-2003 (first entry)

XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 15..16
 FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

XX (LUND/) LUNDQUIST R.

XX (NIEL/) NIELSEN L K.

XX Dziegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 84.7%; Score 127; DB 6; Length 31;
 Best Local Similarity 92.0%; Pred. No. 1.7e-11;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 EASSYDYLWFGGFGGVPEHKKEEN 27
 DB 1 EASSYDYLWFGGFGGVPEHKKEEN 25

RESULT 13

ABR41833
 ID ABR41833 standard; peptide; 31 AA.

XX AC ABR41833;

XX 11-AUG-2003 (first entry)

Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers
 FT Misc-difference 15..16
 FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

XX (LUND/) LUNDQUIST R.

XX (NIEL/) NIELSEN L K.

XX Dziegiel MSH; Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 84.0%; Score 126; DB 6; Length 31;
 Best Local Similarity 92.0%; Pred. No. 2.4e-11;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AKEASSYDYLWFGGFGGVPEHKKE 25

DB 7 AKEASSYDYLWFGGFGGVPEHKKE 31

RESULT 14

ABR41842
 ID ABR41842 standard; peptide; 31 AA.

XX AC ABR41842;

XX 11-AUG-2003 (first entry)

Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 15..16
 FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

XX (LUND/) LUNDQUIST R.

XX (NIEL/) NIELSEN L K.

XX Dziegiel MSH; Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 82.7%; Score 124; DB 6; Length 31;
 Best Local Similarity 91.7%; Pred. No. 4.7e-11;
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 ASSYDYLWFGGFGGVPEHKKEEN 27

DB 1 ASSYDYLWFGGFGGVPEHKKEEN 24

RESULT 15

ABR41832

ID ABR41832 standard; peptide; 31 AA.

XX ABR41832;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:00:42 ; Search time 44.6612 Seconds
(without alignments)
177.141 Million cell updates/sec

Title: US-10-774-602-13

Perfect score: 147

Sequence: 1 PEHKKENMLSHLYVSSKOKENISKEND 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	97.3	28	2 AAR54687	Aar54687 Plasmodiu
2	143	97.3	64	2 AAR54684	Aar54684 Plasmodiu
3	143	97.3	64	6 ABR41813	Abr41813 Merozoite
4	129	87.8	31	6 ABR41855	Abr41855 Peptide u
5	128	87.1	31	6 ABR41854	Abr41854 Peptide u
6	126	85.7	31	6 ABR41853	Abr41853 Peptide u
7	119	81.0	31	6 ABR41852	Abr41852 Peptide u
8	112	76.2	31	6 ABR41851	Abr41851 Peptide u
9	108	73.5	31	6 ABR41850	Abr41850 Peptide u
10	105	71.4	31	6 ABR41849	Abr41849 Peptide u
11	99	67.3	31	6 ABR41848	Abr41848 Peptide u
12	94	63.9	31	6 ABR41847	Abr41847 Peptide u
13	86	58.5	31	6 ABR41846	Abr41846 Peptide u
14	82	55.8	31	6 ABR41845	Abr41845 Peptide u
15	80	54.4	31	6 ABR41844	Abr41844 Peptide u
16	79	53.7	31	6 ABR41843	Abr41843 Peptide u
17	78	53.1	31	6 ABR41842	Abr41842 Peptide u
18	78	53.1	31	6 ABR41841	Abr41841 Peptide u
19	71	48.3	31	6 ABR41840	Abr41840 Peptide u
20	67	45.6	31	6 ABR41839	Abr41839 Peptide u
21	59	40.1	31	6 ABR41838	Abr41838 Peptide u
22	55	37.4	31	6 ABR41837	Abr41837 Peptide u
23	53	36.1	353	4 AAM41451	Aam41451 Human pol
24	53	36.1	426	3 AAB42986	Aab42986 Human ORF
25	53	36.1	426	4 AAM39665	Aam39665 Human pol

26	53	36.1	426	4	AAB94867	Aab94867 Human pro
27	53	36.1	426	4	AAG89197	Aag89197 Human sec
28	53	36.1	426	4	AAG89190	Aag89190 Human sec
29	53	36.1	426	7	ADC06747	Adc06747 Human ino
30	53	36.1	481	4	AAG89295	Aag89295 Human sec
31	53	36.1	3242	4	ABBS8857	Abbs8857 Drosophil
32	52	35.4	443	6	ABU20833	Abu20833 Protein e
33	52	35.4	693	6	ADA14410	Ada14410 Mouse spe
34	52	35.4	2206	3	AAB18254	Aab18254 Plasmodiu
35	51.5	35.0	407	3	AAB18248	Aab18248 Plasmodiu
36	51.5	35.0	1182	6	ABU25521	Abu25521 Protein e
37	51	34.7	31	6	ABR41836	Abr41836 Peptide u
38	51	34.7	174	4	ABBS2591	Abbs2591 Escherich
39	51	34.7	415	6	ABU42518	Abu42518 Protein e
40	50.5	34.4	296	4	ABB69513	Abb69513 Drosophil
41	50	34.0	218	5	ABP08273	Abp08273 Human ORF
42	50	34.0	515	6	ABU25405	Abu25405 Protein e
43	50	34.0	1997	3	AAB18287	Aab18287 Plasmodiu
44	49.5	33.7	774	6	ABM73172	Abm73172 Staphyloc
45	49.5	33.7	1714	3	AAB18275	Aab18275 Plasmodiu

ALIGNMENTS

RESULT 1

AAR54687
ID AAR54687 standard; protein; 28 AA.
AC AAR54687;
XX
XX
DT 25-MAR-2003 (revised)
DT 15-DEC-1994 (first entry)
XX
DE Plasmodium falciparum merozoite 48kD surface protein fragment IV.
XX
XX Merozoite 48kD surface antigen; immunogen; vaccine; malaria;
XX cytophilic antibody.
XX Plasmodium falciparum.
XX
PN WO9409140-A1.
XX
PD 28-APR-1994.
XX
PF 18-OCT-1993; 93WO-FR001024.
XX
PR 19-OCT-1992; 92FR-00012488.
(INSP) INST PASTEUR.
PI Druilhe P, Bouharoun-Tayoun H, Oeuvery C;
XX
XX WPI; 1994-151325/18.
N-PSDB; AAQ64680.
PT New plasmodium antigen generating cytophilic - mimicking protection
induced by long term exposure to the parasite, useful in vaccination
PT against and diagnosis of malaria, and prodn of antibodies for passive
immunisation.
XX
PS Claim 4; Page 37; 52pp; French.
CC This sequence is a preferred fragment of the 48kD merozoite surface
antigen of P.falciparum which is useful as an immunogen to generate
CC antibodies (Abs). The Abs can induce a cytotoxic reaction against the
CC erythrocyte stage of the parasite. (Updated on 25-MAR-2003 to correct PN
field.)
XX
SQ Sequence 28 AA;

Query Match 97.3%; Score 143; DB 2; Length 28;
Best Local Similarity 96.4%; Pred. No. 4.5e-13;

Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
 DB 1 PEHKKEENMLSHLYVSSKDKENISKENE 28

RESULT 2
 AAR54684
 ID AAR54684 standard; protein; 64 AA.
 XX AAR54684;
 AC AAR54684;
 XX 25-MAR-2003 (revised)
 DT 15-DEC-1994 (first entry)
 XX Plasmodium falciparum merozoite 48kD surface protein fragment I.
 DE Merozoite 48kD surface antigen; immunogen; vaccine; malaria;
 XX cytophilic antibody.
 OS Plasmodium falciparum.
 XX WO9409140-A1.
 PN 28-APR-1994.
 PD 18-OCT-1993; 93WO-FR001024.
 XX 19-OCT-1992; 92FR-00012488.
 XX (INSP) INST PASTEUR.
 PA Druiilhe P, Bouharoun-Tayoun H, Oeuvray C;
 PI WPI; 1994-151325/18.
 XX N-PSDB; AAQ64677.
 XX New Plasmodium antigen generating cytophilic - mimicking protection
 PT induced by long term exposure to the parasite, useful in vaccination
 PT against and diagnosis of malaria, and prodn of antibodies for passive
 PT immunisation.
 XX Claim 3; Page 36; 52pp; French.

Query Match 97.1%; Score 143; DB 2; Length 64;
 Best Local Similarity 96.4%; Pred. No. 1.2e-12;
 Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
 DB 37 PEHKKEENMLSHLYVSSKDKENISKENE 64

RESULT 3
 ABR41813
 ID ABR41813 standard; protein; 64 AA.
 XX ABR41813;
 AC ABR41813;
 XX 11-AUG-2003 (first entry)
 DT Merozoite surface antigen MSP-3 peptide (aa194-257).
 DE

KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 immunotherapy; vaccine.
 XX Plasmodium falciparum.
 OS Key Location/Qualifiers
 FH Region 27. .37
 FT /note= "epitope specifically referred to in Claim 3"
 XX GB2378949-A.
 PN 26-FEB-2003.
 XX 16-AUG-2001; 2001GB-00020057.
 PF 16-AUG-2001; 2001GB-00020057.
 PR (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX Dziegiel MSH, Lundquist R, Nielsen LK;
 PI WPI; 2003-303108/30.
 DR Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX Disclosure; Fig 3; 75pp; English.
 PS The present sequence is that of amino acids 194-257 of the merozoite
 CC surface protein 3 (MSP-3) of Plasmodium falciparum clone D10. The
 CC invention relates to the construction of recombinant human antibodies
 CC against MSP-3, in particular MSP-3(194-257) antigen. The antibodies (see
 CC also ABR41816-21) are able to passively induce naturally occurring non-
 CC sterile malaria immunity (termed premunition) and exert their effect via
 CC the antibody-dependent cellular inhibition of growth (ADCI) mechanism.
 CC They can be used in the diagnosis and treatment of malaria. The invention
 CC also provides a method of screening a phage library with target
 CC molecules, such as MSP-3 antigen, bound to beads, preferably magnetic
 CC microspheres
 XX Sequence 64 AA;
 SQ Query Match 97.3%; Score 143; DB 6; Length 64;
 Best Local Similarity 96.4%; Pred. No. 1.2e-12;
 Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
 DB 37 PEHKKEENMLSHLYVSSKDKENISKENE 64

RESULT 4
 ABR41855
 ID ABR41855 standard; peptide; 31 AA.
 XX ABR41855;
 AC ABR41855;
 DT 11-AUG-2003 (first entry)
 XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 DE Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX Plasmodium falciparum.
 OS Key Location/Qualifiers
 FH Misc-difference 15. .16
 FT /note= "linker residues"
 FT

GN GB2378949-A.
 XX PD 26-FEB-2003.
 XX PF 16-AUG-2001; 2001GB-00020057.
 XX PR 16-AUG-2001; 2001GB-00020057.
 XX PA (DZIE/) DZIEGIEL M S H.
 XX PA (LUND/) LUNDQUIST R.
 XX PA (NIEL/) NIELSEN L K.
 XX PI Driegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX Example 3; Fig 10; 75pp; English.
 CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX SQ Sequence 31 AA;
 Query Match 87.8%; Score 129; DB 6; Length 31;
 Best Local Similarity 89.3%; Pred. No. 4.8e-11;
 Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 PEHKKENMLSHLYVSSKDKENISKEND 28
 DB 4 PEHKKENMLSHLYVSSKDKENISKENE 31
 ID 1854
 AC ABR41854 standard; peptide; 31 AA.
 XX ABR41854;
 DT 11-AUG-2003 (first entry)
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX Plasmodium falciparum.
 XX Key Location/Qualifiers
 FT Misc-difference 15..16
 FT /note= "linker residues"
 XX GB2378949-A.
 XX PD 26-FEB-2003.
 XX PF 16-AUG-2001; 2001GB-00020057.
 XX PR 16-AUG-2001; 2001GB-00020057.
 XX PA (DZIE/) DZIEGIEL M S H.
 XX PA (LUND/) LUNDQUIST R.
 XX PI Driegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX Example 3; Fig 10; 75pp; English.
 CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX SQ Sequence 31 AA;
 Query Match 87.8%; Score 129; DB 6; Length 31;
 Best Local Similarity 89.3%; Pred. No. 4.8e-11;
 Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 PEHKKENMLSHLYVSSKDKENISKEND 28
 DB 4 PEHKKENMLSHLYVSSKDKENISKENE 31
 ID 1854
 AC ABR41854 standard; peptide; 31 AA.
 XX ABR41854;
 DT 11-AUG-2003 (first entry)
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX Plasmodium falciparum.
 XX Key Location/Qualifiers
 FT Misc-difference 15..16
 FT /note= "linker residues"
 XX GB2378949-A.
 XX PD 26-FEB-2003.
 XX PF 16-AUG-2001; 2001GB-00020057.
 XX PR 16-AUG-2001; 2001GB-00020057.
 XX PA (DZIE/) DZIEGIEL M S H.
 XX PA (LUND/) LUNDQUIST R.

XX 16-AUG-2001; 2001GB-00020057.
 XX PR (DZIE/) DZIEGIEL M S H.
 XX PA (LUND/) LUNDQUIST R.
 XX PA (NIEL/) NIELSEN L K.
 XX PI Driegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX Example 3; Fig 10; 75pp; English.
 CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX SQ Sequence 31 AA;
 Query Match 87.1%; Score 128; DB 6; Length 31;
 Best Local Similarity 92.6%; Pred. No. 6.6e-11;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 PEHKKENMLSHLYVSSKDKENISKEN 27
 DB 5 PEHKKENMLSHLYVSSKDKENISKEN 31
 RESULT 6
 ABR41853
 ID ABR41853 standard; peptide; 31 AA.
 XX ABR41853;
 AC ABR41853;
 DT 11-AUG-2003 (first entry)
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX Plasmodium falciparum.
 XX Key Location/Qualifiers
 FT Misc-difference 15..16
 FT /note= "linker residues"
 XX GB2378949-A.
 XX PD 26-FEB-2003.
 XX PF 16-AUG-2001; 2001GB-00020057.
 XX PR 16-AUG-2001; 2001GB-00020057.
 XX PA (DZIE/) DZIEGIEL M S H.
 XX PA (LUND/) LUNDQUIST R.

PS Example 3; Fig 10; 75pp; English.

CC The present sequence is one of a set of 34 peptides (see ABR41822-55) used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues 15 and 16 replaced by proprietary linker residues. The peptides were examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are encoded by clones obtained from a phage display system based on antibody variable region genes from the peripheral blood leucocytes of malaria immune human individuals. RAM2 reacted with the N-terminal, middle and C-terminal part of the MSP-3 antigen, with highest reactivity with the middle part. RAM1 reacted only weakly with the peptides. The invention provides recombinant human antibodies against MSP-3, which are useful for the diagnosis and treatment of malaria. The recombinant human antibodies include the complementarity determining regions from the light and heavy chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 76.2%; Score 112; DB 6; Length 31;

Best Local Similarity 91.7%; Pred. No. 1.2e-08;

Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PEHKEENMLSHLYVSSKDKENIS 24

Db 8 PEHKEEXXLHLYVSSKDKENIS 31

RESULT 9

ABR41850

ID ABR41850 standard; peptide; 31 AA.

XX AC ABR41850;

DT 11-AUG-2003 (first entry)

DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;

KW immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 15..16

FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

PF 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

PA (LUND/) LUNDQUIST R.

PA (NIEL/) NIELSEN L K.

XX Driegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3 antigen of Plasmodium falciparum, useful for treatment and prophylaxis of a malarial disease.

PS Example 3; Fig 10; 75pp; English.

CC The present sequence is one of a set of 34 peptides (see ABR41822-55) used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino

CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues 15 and 16 replaced by proprietary linker residues. The peptides were examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are encoded by clones obtained from a phage display system based on antibody variable region genes from the peripheral blood leucocytes of malaria immune human individuals. RAM2 reacted with the N-terminal, middle and C-terminal part of the MSP-3 antigen, with highest reactivity with the middle part. RAM1 reacted only weakly with the peptides. The invention provides recombinant human antibodies against MSP-3, which are useful for the diagnosis and treatment of malaria. The recombinant human antibodies include the complementarity determining regions from the light and heavy chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 73.5%; Score 108; DB 6; Length 31;

Best Local Similarity 91.3%; Pred. No. 4.3e-08;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PEHKEENMLSHLYVSSKDKENI 23

Db 9 PEHKEEXXLHLYVSSKDKENI 31

RESULT 10

ABR41849

ID ABR41849 standard; peptide; 31 AA.

XX AC ABR41849;

DT 11-AUG-2003 (first entry)

DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;

KW immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 15..16

FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

PF 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

PA (LUND/) LUNDQUIST R.

PA (NIEL/) NIELSEN L K.

XX Driegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3 antigen of Plasmodium falciparum, useful for treatment and prophylaxis of a malarial disease.

PS Example 3; Fig 10; 75pp; English.

CC The present sequence is one of a set of 34 peptides (see ABR41822-55) used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues 15 and 16 replaced by proprietary linker residues. The peptides were examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are encoded by clones obtained from a phage display system based on antibody variable region genes from the peripheral blood leucocytes of malaria

CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 71.4%; Score 105; DB 6; Length 31;
 Best Local Similarity 90.9%; Pred. NO. 1.1e-07;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKOKEN 22
 DB 10 PEHKXXENMLSHLYVSSKOKEN 31

DT 11

ABR41848
 ID ABR41848 standard; peptide; 31 AA.

AC ABR41848;

DT 11-AUG-2003 (first entry)

DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.

KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.

OS Plasmodium falciparum.

FH Key Location/Qualifiers
 FT Misc-difference 15..16 /note= "linker residues"

XX GB2378949-A.

PD 26-FEB-2003.

PF 16-AUG-2001; 2001GB-00020057.

PR 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.
 XX (LUND/) LUNDQUIST R.
 XX (NIEL/) NIELSEN L K.

PI Dzigiel MSH, Lundquist R, Nielsen LK;
 WPI; 2003-303108/30.

PT Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

PS Example 3; Fig 10; 75pp; English.

CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies

CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 67.3%; Score 99; DB 6; Length 31;
 Best Local Similarity 90.5%; Pred. No. 8e-07;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDXE 21
 DB 11 PEHKXXENMLSHLYVSSKDXE 31

RESULT 12

ABR41847

ID ABR41847 standard; peptide; 31 AA.

AC ABR41847;

DT 11-AUG-2003 (first entry)

DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.

KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.

OS Plasmodium falciparum.

FH Key Location/Qualifiers
 FT Misc-difference 15..16 /note= "linker residues"

XX GB2378949-A.

PD 26-FEB-2003.

PF 16-AUG-2001; 2001GB-00020057.

PR 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.
 XX (LUND/) LUNDQUIST R.
 XX (NIEL/) NIELSEN L K.

PI Dzigiel MSH, Lundquist R, Nielsen LK;
 WPI; 2003-303108/30.

PT Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

PS Example 3; Fig 10; 75pp; English.

CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 63.9%; Score 94; DB 6; Length 31;
 Best Local Similarity 90.0%; Pred. No. 4.1e-06;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PEKKKEENMLSHLYVSSKD 20
 DB 12 PEKKKEENMLSHLYVSSKD 31

RESULT 13
 ABR41846
 ID ABR41846 standard; peptide; 31 AA.
 AC ABR41846;
 DT 11-AUG-2003 (first entry)
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 XX immunotherapy; vaccine; epitope.
 XX Plasmodium falciparum.

Key Location/Qualifiers
 Misc-difference 15. .16
 /note= "linker residues"

GB2378949-A.
 26-FEB-2003.
 16-AUG-2001; 2001GB-00020057.
 16-AUG-2001; 2001GB-00020057.
 (DZIE/) DZIEGIEL M S H.
 (LUND/) LUNDQUIST R.
 (NIEL/) NIELSEN L K.
 Driegiel MSH, Lundquist R, Nielsen LK;
 WPI; 2003-303108/30.

Novel recombinant human antibody specific for merozoite surface protein-3 antigen of Plasmodium falciparum, useful for treatment and prophylaxis of a malarial disease.

Example 3; Fig 10; 75pp; English.

The present sequence is one of a set of 34 peptides (see ABR41822-55) used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues 15 and 16 replaced by proprietary linker residues. The peptides were examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are encoded by clones obtained from a phage display system based on antibody variable region genes from the peripheral blood leucocytes of malaria immune human individuals. RAM2 reacted with the N-terminal, middle and C-terminal part of the MSP-3 antigen, with highest reactivity with the middle part. RAM1 reacted only weakly with the peptides. The invention provides recombinant human antibodies against MSP-3, which are useful for the diagnosis and treatment of malaria. The recombinant human antibodies include the complementarity determining regions from the light and heavy chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

Sequence 31 AA;
 Query Match 58.5%; Score 86; DB 6; Length 31;
 Best Local Similarity 89.5%; Pred. No. 5.4e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PEKKKEENMLSHLYVSSKD 19

Query Match 55.8%; Score 82; DB 6; Length 31;
 Best Local Similarity 100.0%; Pred. No. 0.0002;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PEKKKEENMLSHLYV 15
 DB 17 PEKKKEENMLSHLYV 31

RESULT 15

ABR41845
ID ABR41845 standard; peptide; 31 AA.
XX
AC ABR41845;
XX
DT 11-AUG-2003 (first entry)
XX
DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
XX
KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
KW immunotherapy; vaccine; epitope.
XX
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FT Misc-difference 15..16
FT /note= "linker residues"
XX
GB2378949-A.
26-FEB-2003.
XX
PF 16-AUG-2001; 2001GB-00020057.
XX
PR 16-AUG-2001; 2001GB-00020057.
XX
PA (DZIE/) DZIEGIEL M S H.
PA (LUND/) LUNDQUIST R.
PA (NIEL/) NIELSEN L K.
XX
PI Dziegiel MSH, Lundquist R, Nielsen LK;
XX
DR WPI; 2003-303108/30.
XX
PT Novel recombinant human antibody specific for merozoite surface protein-3
PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
PT a malarial disease.
XX
PS Example 3; Fig 10; 75pp; English.
XX
CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
CC used in a pepSCAN epitope mapping of Plasmodium falciparum merozoite
CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
CC 15 and 16 replaced by proprietary linker residues. The peptides were
CC examined for reactivity with RAM1 and RAM2 IGI1. RAM1 and RAM2 are
CC encoded by clones obtained from a phage display system based on antibody
CC variable region genes from the peripheral blood leucocytes of malaria
CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
CC terminal part of the MSP-3 antigen, with highest reactivity with the
CC middle part. RAM1 reacted only weakly with the peptides. The invention
CC provides recombinant human antibodies against MSP-3, which are useful for
CC the diagnosis and treatment of malaria. The recombinant human antibodies
CC include the complementarity determining regions from the light and heavy
CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
SQ Sequence 31 AA;
Query Match 54.4%; Score 80; DB 6; Length 31;
Best Local Similarity 88.9%; Pred. No. 0.00038;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 PEHKKENMLSHLYVSK 18
Db 14 PXXKKEENMLSHLYVSK 31

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:47 ; Search time 26.4463 Seconds
(without alignments)
298.263 Million cell updates/sec

Title: US-10-774-602-11
Perfect score: 124
Sequence: 1 YEKAKYQKANOAVLKAKEASSYD 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	124	100.0	354	5	Q25995	Q25995 plasmodium
2	124	100.0	354	5	Q81355	Q81355 plasmodium
3	124	100.0	380	5	Q26019	Q26019 plasmodium
4	121	97.6	361	5	Q95815	Q95815 plasmodium
5	121	97.6	379	5	Q25705	Q25705 plasmodium
6	121	97.6	379	5	Q25706	Q25706 plasmodium
7	121	97.6	379	5	Q9U6C4	Q9U6C4 plasmodium
8	120	96.8	329	5	Q9NFV9	Q9NFV9 plasmodium
9	119	96.0	346	5	Q9U0G0	Q9U0G0 plasmodium
10	118	95.2	157	5	Q15854	Q15854 plasmodium
11	118	95.2	157	5	Q15855	Q15855 plasmodium
12	118	95.2	157	5	Q9TV90	Q9TV90 plasmodium
13	118	95.2	157	5	Q15856	Q15856 plasmodium
14	118	95.2	157	5	Q15853	Q15853 plasmodium
15	118	95.2	157	5	Q9TV91	Q9TV91 plasmodium
16	116	93.5	64	5	Q15762	Q15762 plasmodium

17	115	92.7	183	5	Q9TY89	Q9TY89 plasmodium
18	115	92.7	183	5	Q15857	Q15857 plasmodium
19	115	92.7	183	5	Q9TV87	Q9TV87 plasmodium
20	115	92.7	183	5	Q15858	Q15858 plasmodium
21	115	92.7	183	5	Q9TY88	Q9TY88 plasmodium
22	115	92.7	183	5	Q9TY86	Q9TY86 plasmodium
23	115	92.7	183	5	Q15859	Q15859 plasmodium
24	115	92.7	183	5	Q15862	Q15862 plasmodium
25	111	89.5	190	5	Q15860	Q15860 plasmodium
26	110	88.7	183	5	Q15863	Q15863 plasmodium
27	108	87.1	180	5	Q9NTG7	Q9NTG7 plasmodium
28	102	82.3	183	5	Q15861	Q15861 plasmodium
29	59	47.6	890	2	Q01891	Q01891 enterococcu
30	59	47.6	890	16	Q82YS1	Q82YS1 enterococcu
31	57	46.0	639	5	Q86RM4	Q86RM4 plasmodium
32	55.5	44.8	652	10	Q9ASS0	Q9ASS0 arabidopsis
33	55.5	44.8	820	10	Q9LFE0	Q9LFE0 arabidopsis
34	55	44.4	752	16	Q9CGC9	Q9CGC9 lactococcu
35	55	44.4	891	16	Q04111	Q04111 enterococcu
36	54.5	44.0	1037	3	Q9P3K1	Q9P3K1 neutrospora
37	52	41.9	213	5	P90899	P90899 caenorhabdi
38	52	41.9	231	5	Q86D10	Q86D10 caenorhabdi
39	52	41.9	420	10	Q9LZL9	Q9LZL9 arabidopsis
40	51.5	41.5	647	10	Q7X7G6	Q7X7G6 oryza sativ
41	51	41.1	169	5	Q9NIG6	Q9NIG6 plasmodium
42	51	41.1	420	10	Q8LFK4	Q8LFK4 arabidopsis
43	51	41.1	742	2	O50582	O50582 streptococ
44	51	41.1	901	16	Q839L8	Q839L8 enterococcu
45	51	41.1	972	2	Q8KYV2	Q8KYV2 uncultured

ALIGNMENTS

RESULT 1

Q25995	ID	Q25995	PRELIMINARY;	PRT;	354 AA.
AC	Q25995;				
DT	01-NOV-1996	(T-EMBLrel. 01, Created)			
DT	01-NOV-1996	(T-EMBLrel. 01, Last sequence update)			
DT	01-OCT-2002	(T-EMBLrel. 22, Last annotation update)			
DE	Antigen.				
OS	Plasmodium falciparum (isolate NF54).				
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.				
OX	NCBI_TaxID=5843;				
RN	[1]	TaxID=5843;			
RP	SEQUENCE FROM N.A.				
RC	STRAIN=NF54;				
RX	MEDLINE=95198774; PubMed=7891748;				
RA	McColl D.J., Silva A., Foley M., Kun J.F., Favaloro J.M., Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;				
RA	"Molecular variation in a novel polymorphic antigen associated with				
RT	Plasmodium falciparum merozoites.";				
RL	Mol. Biochem. Parasitol. 68:53-67(1994).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=NF54;				
RX	MEDLINE=98156743; PubMed=9497029;				
RA	McColl D.J., Anders R.F.;				
RT	"Conservation of structural motifs and antigenic diversity in the				
RT	Plasmodium falciparum merozoite surface protein-3 (MSP-3).";				
RL	Mol. Biochem. Parasitol. 90:21-31(1997).				
DR	EMBL; L28825; AAC09377.1;				
SQ	SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;				

Query Match 100.0%; Score 124; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 5, 1e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKYQKANOAVLKAKEASSYD 25
DB 167 YEKAKYQKANOAVLKAKEASSYD 191

```
RESULT 2
Q8IJ55 PRELIMINARY; PRT; 354 AA.
ID Q8IJ55
AC Q8IJ55
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Merozoite surface protein 3.
GN PF10_0345.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
  Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
  Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
  Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
  Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
  Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
  McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
  Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
  Fraser C.M., Barrell B.,
  RA "Genome sequence of the human malaria parasite Plasmodium
  falciparum."
RT Nature 419:498-511 (2002).
RL EMBL; AE014834; AAN35542.1; -.
SQ SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;

Query Match 100.0%; Score 124; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.1e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
Db 167 YEKAKNAYOKANQAVLKAKEASSYD 191

RESULT 3
Q26019 PRELIMINARY; PRT; 380 AA.
ID Q26019
AC Q26019
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polymorphic antigen precursor.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC27;
RX MEDLINE=95198774; PubMed=7891748;
RA McCall D.J., Silva A., Foley M., Kun J.F., Favaloro J.M.,
  Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;
RT "Molecular variation in a novel polymorphic antigen associated with
  Plasmodium falciparum merozoites."
RL Mol. Biochem. Parasitol. 68:53-67 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FC27;
RX MEDLINE=98156743; PubMed=9497029;
RA McCall D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
  Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31 (1997).
DR EMBL; L07944; AAC09378.1; -.
KW SIGNAL.
FT SIGNAL 1 25 POTENTIAL.

FT CHAIN 26 380 POLYMORPHIC ANTIGEN.
SQ SEQUENCE 380 AA; 43290 MW; 0986CA1393094CA2 CRC64;

Query Match 100.0%; Score 124; DB 5; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.4e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
Db 194 YEKAKNAYOKANQAVLKAKEASSYD 218

RESULT 4
Q95P15 PRELIMINARY; PRT; 361 AA.
ID Q95P15
AC Q95P15
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVO;
RA Hisaeda H., Saul A., Long C.A., Miller L.H., Stowers A.W.;
RT "Merozoite Surface Protein 3 and Protection Against Malaria in Aotus
  non-ter monkeys."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY044180; AAK94780.1; -.
FT NON_TER 1
FT NON_TER 361
SQ SEQUENCE 361 AA; 41163 MW; 6127A3041587BA74 CRC64;

Query Match 97.6%; Score 121; DB 5; Length 361;
Best Local Similarity 96.0%; Pred. No. 1.3e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
Db 173 YERAKNAYOKANQAVLKAKEASSYD 197

RESULT 5
Q25705 PRELIMINARY; PRT; 379 AA.
ID Q25705
AC Q25705
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X1;
RX MEDLINE=98156743; PubMed=9497029;
RA McCall D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
  Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31 (1997).
DR EMBL; U08851; AAC47831.1; -.
SQ SEQUENCE 379 AA; 43344 MW; DC7AF106887C9AA0 CRC64;

Query Match 97.6%; Score 121; DB 5; Length 379;
Best Local Similarity 96.0%; Pred. No. 1.4e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
Db 193 YERAKNAYOKANQAVLKAKEASSYD 217
```

```
RESULT 6
Q25706 PRELIMINARY; PRT; 379 AA.
AC Q25706;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2002 (TrEMBLrel. 01, Last sequence update)
DE Polymorphic antigen.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=csl2;
RX MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
RT Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; U08852; AAC47832.1; -.
SQ SEQUENCE 379 AA; 43302 MW; ABF9D54E1ED91A24 CRC64;

Query Match 97.6%; Score 121; DB 5; Length 379;
Best Local Similarity 96.0%; Pred. No. 1.4e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKANOAVLKAKEASSYD 25
Db 193 YERAKNAYQKANOAVLKAKEASSYD 217

RESULT 7
Q9U6C4 PRELIMINARY; PRT; 379 AA.
AC Q9U6C4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
GN MSP-3.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC1/HN;
RX Li X.R., Yu X.B., Luo S.H., Shan Z.X., Fang J.M., Xu J.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF188190; AAF04099.1; -.
SQ SEQUENCE 379 AA; 43316 MW; C152A54E1F9D5F25 CRC64;

Query Match 97.6%; Score 121; DB 5; Length 379;
Best Local Similarity 96.0%; Pred. No. 1.4e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKANOAVLKAKEASSYD 25
Db 193 YERAKNAYQKANOAVLKAKEASSYD 217

RESULT 8
Q9NFV9 PRELIMINARY; PRT; 329 AA.
AC Q9NFV9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (fragment).
GN MSP3.
```

```
OS Plasmodium falciparum (isolate 7G8).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7G8;
RX MEDLINE=20416497; PubMed=10960178;
RA Okenu D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in
RT Plasmodium reichenowi and Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 109:185-188(2000).
DR EMBL; AJ252287; CAB85901.1; -.
FT NON_TER 1
FT NON_TER 329
SQ SEQUENCE 329 AA; 36916 MW; C5B045DB5E21A159 CRC64;

Query Match 96.8%; Score 120; DB 5; Length 329;
Best Local Similarity 96.0%; Pred. No. 1.7e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKANOAVLKAKEASSYD 25
Db 157 YERAKNAYQKANOAVLKAKEASSYD 181

RESULT 9
Q9U0G0 PRELIMINARY; PRT; 346 AA.
AC Q9U0G0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (fragment).
GN MSP3.
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20416497; PubMed=10960178;
RA Okenu D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in
RT Plasmodium reichenowi and Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 109:185-188(2000).
DR EMBL; AJ252286; CAB65754.1; -.
FT NON_TER 1
FT NON_TER 346
SQ SEQUENCE 346 AA; 39127 MW; A804B96BDFAF010 CRC64;

Query Match 96.0%; Score 119; DB 5; Length 346;
Best Local Similarity 96.0%; Pred. No. 2.4e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKANOAVLKAKEASSYD 25
Db 170 YERAKNAYQKANOAVLKAKEASSYD 194

RESULT 10
O15854 PRELIMINARY; PRT; 157 AA.
AC O15854;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Merozoite surface protein 3 (fragment).
GN SPAM.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97391128; PubMed=9247935;
```

RA Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RL surface protein 3.";
RL Mol. Biochem. Parasitol. 87:231-234 (1997).
DR EMBL; AF001138; AAC47663.1; -.
FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17009 MW; PDF17D10C45E74DB CRC64;

Query Match 95.2%; Score 118; DB 5; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKQANQAVLKAKEASSY 24
|||
Db 134 YEKAKNAYQKQANQAVLKAKEASSY 157

T 11
O 5
ID O15855 PRELIMINARY; PRT; 157 AA.
AC O15855;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN SPAM.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97391128; PubMed=9247935;
RA Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RL surface protein 3.";
RL Mol. Biochem. Parasitol. 87:231-234 (1997).
DR EMBL; AF001140; AAC47665.1; -.
FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 16981 MW; E2F6D10C44963DB CRC64;

Query Match 95.2%; Score 118; DB 5; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKQANQAVLKAKEASSY 24
|||
Db 134 YEKAKNAYQKQANQAVLKAKEASSY 157

RESULT 12
Q9TY90
ID Q9TY90 PRELIMINARY; PRT; 157 AA.
AC Q9TY90;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN SPAM.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97391128; PubMed=9247935;
RA Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RL surface antigen 3.";
RL Mol. Biochem. Parasitol. 87:231-234 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=01/3;

RA Eisen D.P., Saul A., Fryauff D.J., Reeder J.C., Coppel R.L.;
RT "Extended antigenic haplotyping of Plasmodium falciparum field
RT isolates shows marked complexity.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF001142; AAC47667.1; -.
DR EMBL; AF213689; AAF59916.1; -.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17007 MW; F06E6080DB4963DB CRC64;

Query Match 95.2%; Score 118; DB 5; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKQANQAVLKAKEASSY 24
|||
Db 134 YEKAKNAYQKQANQAVLKAKEASSY 157

RESULT 13
O15856
ID O15856 PRELIMINARY; PRT; 157 AA.
AC O15856;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN SPAM.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97391128; PubMed=9247935;
RA Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RL surface protein 3.";
RL Mol. Biochem. Parasitol. 87:231-234 (1997).
DR EMBL; AF001141; AAC47666.1; -.
FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17034 MW; F12B7B817053D270 CRC64;

Query Match 95.2%; Score 118; DB 5; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKQANQAVLKAKEASSY 24
|||
Db 134 YEKAKNAYQKQANQAVLKAKEASSY 157

RESULT 14
O15853
ID O15853 PRELIMINARY; PRT; 157 AA.
AC O15853;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN SPAM.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97391128; PubMed=9247935;
RA Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RL surface protein 3.";
RL Mol. Biochem. Parasitol. 87:231-234 (1997).
DR EMBL; AF001137; AAC47662.1; -.
RN [2]


```

FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 16981 MW; E2FE6D10C4963DB CRC64;

Query Match
Best Local Similarity 95.2%; Score 118; DB 5; Length 157;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YEKAKNAYQKANOAVLKAKEASSY 24
Db 134 YEKAKNAYQKANOAVLKAKEASSY 157

RESULT 15
Q9TY91
ID Q9TY91 PRELIMINARY; PRT; 157 AA.
AC Q9TY91;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Merozoite surface protein 3 (Fragment).
SPAM.
OC Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5633;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97391128; PubMed=9247935;
RA Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RL surface protein 3."
RL Mol. Biochem. Parasitol. 87:231-234 (1997).
DR EMBL; AF001139; AAC47664.1; -.
FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17007 MW; F06E6080DB4963DB CRC64;

Query Match
Best Local Similarity 95.2%; Score 118; DB 5; Length 157;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YEKAKNAYQKANOAVLKAKEASSY 24
Db 134 YEKAKNAYQKANOAVLKAKEASSY 157

```

Search completed: September 22, 2004, 18:09:05
 Job time : 29.4463 secs

100-100000-100000

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:47 ; Search time 28.562 Seconds
(without alignments)
298.263 Million cell updates/sec

Title: US-10-774-602-12
Perfect score: 150
Sequence: 1 AKASDYVILGWFGGVPEHKKEEN 27

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

1 number of hits satisfying chosen parameters: 1017041

Maximum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phase: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_rvirus: *
16: sp_bacteriaph: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	150	100.0	64	5	015762 plasmodium
2	150	100.0	329	5	Q9NFV9
3	150	100.0	354	5	Q25995 plasmodium
4	150	100.0	354	5	Q8IJ55 plasmodium
5	150	100.0	361	5	Q95P15 plasmodium
6	150	100.0	379	5	Q25705 plasmodium
7	150	100.0	379	5	Q25706 plasmodium
8	150	100.0	379	5	Q9U6C4 plasmodium
9	150	100.0	380	5	Q26019 plasmodium
10	145	96.7	346	5	Q9U0G0 plasmodium
11	66.5	44.3	371	5	Q9GS22 plasmodium
12	66.5	44.3	371	5	Q8IJ54 plasmodium
13	66	44.0	405	5	Q8IJ48 plasmodium
14	61	40.7	1238	16	O69947 streptomyces
15	58	38.7	596	16	Q7WR42 bordetella
16	58	38.7	596	16	Q7W275 bordetella

```

17 38.7 596 16 Q7W001
18 38.0 382 16 Q7UDT7
19 38.0 391 16 Q83MN5
20 36.7 479 16 Q25715
21 36.7 479 16 Q9ZM80
22 36.0 533 9 Q94MX6
23 35.3 591 16 Q8EP50
24 35.3 682 10 Q9M3E5
25 34.7 676 10 Q8H411
26 34.7 858 5 Q17647
27 34.7 860 5 Q95NM4
28 34.0 283 16 Q8ZMG6
29 34.0 271 16 O50860
30 34.0 698 10 Q7XWV8
31 33.3 389 10 Q8GZY3
32 33.3 433 10 Q84SN3
33 33.3 598 16 Q88VQ8
34 33.3 599 16 Q8EP55
35 33.3 647 10 Q8H413
36 33.3 701 10 Q852A9
37 33.3 735 10 Q9SW88
38 33.0 168 10 Q22625
39 33.0 848 16 Q82WX2
40 33.0 2021 5 O62191
41 32.7 197 12 P89214
42 32.7 382 16 Q7U6Q2
43 32.7 561 5 Q8IJ58
44 32.7 587 10 Q9LYU1
45 32.7 613 10 Q9M1G4

```

ALIGNMENTS

RESULT 1

```

015762 PRELIMINARY; PRT; 64 AA.
ID O15762
AC O15762;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN MSP-3.
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NF54;
RX MEDLINE=94348085; PubMed=8068948;
RA Oauvray C., Bouharoun-Tayoun H., Gras-Masse H., Bottius E., Kaidoh T.,
RA Aikawa M., Filgueira M.-C., Tartar A., Drulhe P.;
RT "Merozoite surface protein-3: a malaria protein inducing antibodies
RT that promote Plasmodium falciparum killing by cooperation with blood
RT monocytes."
RL Blood 84:1594-1602 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NF54;
RA Oauvray C., Bouharoun-Tayoun H., Gras-Masse H., Bottius E., Kaidoh T.,
RA Aikawa M., Filgueira M.-C., Tartar A., Drulhe P.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF024624; AAB81610.1; -.
FT NON_TER
FT NON_TER
SQ SEQUENCE 64 AA; 7326 MW; 02E0ABF614C15A3B CRC64;

```

Query Match 100.0%; Score 150; DB 5; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKASDYVILGWFGGVPEHKKEEN 27

```
Db 18 AKEASSYDYLWGFEGGVPEHKKEEN 44
RESULT 2
Q9NFV9 ID Q9NFV9 PRELIMINARY; PRT; 329 AA.
AC Q9NFV9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN MSP3.
OS Plasmodium falciparum (isolate 7G8).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7G8;
RX MEDLINE=20416497; PubMed=10960178;
RA Okenu D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in Plasmodium reichenowi and Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 109:185-188 (2000).
DR EMBL; AJ252287; CAB85901.1; -.
FT NON_TER 1
FT NON_TER 329
SQ SEQUENCE 329 AA; 36916 MW; C5B045DB5E21A159 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLWGFEGGVPEHKKEEN 27
Db 174 AKEASSYDYLWGFEGGVPEHKKEEN 200
RESULT 3
Q25995 ID Q25995 PRELIMINARY; PRT; 354 AA.
AC Q25995;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Antigen.
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NF54;
RX MEDLINE=95198774; PubMed=7891748;
RA McColl D.J., Silva A., Foley M., Kun J.F., Favalaro J.M.,
RA Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;
RT "Molecular variation in a novel polymorphic antigen associated with Plasmodium falciparum merozoites.";
RL Mol. Biochem. Parasitol. 68:53-67 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NF54;
RX MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31 (1997).
DR EMBL; L28825; AAC09377.1; -.
SQ SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLWGFEGGVPEHKKEEN 27
Db 184 AKEASSYDYLWGFEGGVPEHKKEEN 210
RESULT 4
Q8IJ55 ID Q8IJ55 PRELIMINARY; PRT; 354 AA.
AC Q8IJ55;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Merozoite surface protein 3.
GN PF10_0345.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium falciparum.";
RL Nature 419:498-511 (2002).
DR EMBL; AE014834; RAN35542.1; -.
SQ SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLWGFEGGVPEHKKEEN 27
Db 184 AKEASSYDYLWGFEGGVPEHKKEEN 210
RESULT 5
Q95P15 ID Q95P15 PRELIMINARY; PRT; 361 AA.
AC Q95P15;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EVO;
RA Hissaeda H., Saul A., Long C.A., Miller L.H., Stowers A.W.;
RT "Merozoite Surface Protein 3 and Protection Against Malaria in Aotus RT Monkeys.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY04180; AAK94780.1; -.
FT NON_TER 1
FT NON_TER 361
SQ SEQUENCE 361 AA; 41163 MW; 6127A3041587BA74 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLWGFEGGVPEHKKEEN 27
```

```
Db 190 AKEASSYDYLIGWFGGGVPEHKKEEN 216

RESULT 6
Q25705 PRELIMINARY; PRT; 379 AA.
ID Q25705
AC Q25705
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RP STRAIN=FC27;
RC MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
   Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; U08851; AAC47831.1; -.
SQ SEQUENCE 379 AA; 43344 MW; DC7AF106887C8AA0 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLIGWFGGGVPEHKKEEN 27
Db 210 AKEASSYDYLIGWFGGGVPEHKKEEN 236

RESULT 7
Q25706 PRELIMINARY; PRT; 379 AA.
ID Q25706
AC Q25706
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RP STRAIN=FC27;
RC MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
   Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; U08852; AAC47832.1; -.
SQ SEQUENCE 379 AA; 43302 MW; ABF9D54E1ED91A24 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLIGWFGGGVPEHKKEEN 27
Db 210 AKEASSYDYLIGWFGGGVPEHKKEEN 236

RESULT 8
Q9U6C4 PRELIMINARY; PRT; 379 AA.
ID Q9U6C4
AC Q9U6C4
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RP STRAIN=FC27;
RC MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
   Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; U08853; AAC47833.1; -.
SQ SEQUENCE 379 AA; 43302 MW; ABF9D54E1ED91A24 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLIGWFGGGVPEHKKEEN 27
Db 210 AKEASSYDYLIGWFGGGVPEHKKEEN 236

RESULT 9
Q26019 PRELIMINARY; PRT; 380 AA.
ID Q26019
AC Q26019
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polymorphic antigen precursor.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RP STRAIN=FC27;
RC MEDLINE=98156743; PubMed=7891748;
RA McColl D.J., Silva A., Foley M., Kun J.F., Favaloro J.M.,
   Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;
RT "Molecular variation in a novel polymorphic antigen associated with
   Plasmodium falciparum merozoites.";
RL Mol. Biochem. Parasitol. 68:53-67(1994).
DR EMBL; L07944; AAC09378.1; -.
SQ SEQUENCE 380 AA; 43290 MW; 0986CA1393094CA2 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLIGWFGGGVPEHKKEEN 27
Db 211 AKEASSYDYLIGWFGGGVPEHKKEEN 237

RESULT 10
Q9U0G0 PRELIMINARY; PRT; 346 AA.
ID Q9U0G0
AC Q9U0G0
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Polymorphic antigen.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RP STRAIN=FC27;
RC MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
   Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; L07944; AAC09378.1; -.
SQ SEQUENCE 380 AA; 43290 MW; 0986CA1393094CA2 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLIGWFGGGVPEHKKEEN 27
Db 211 AKEASSYDYLIGWFGGGVPEHKKEEN 237
```

```

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN MSP3.
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20416497; PubMed=10960178;
RA Okeni D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in
RT Plasmodium reichenowi and Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 109:185-188(2000).
DR EMBL; AJ252286; CAB65754.1; -.
FT NON_TER 1
FT NON_TER 346
SQ SEQUENCE 346 AA; 39127 MW; A804B96BDFAF010 CRC64;

Query Match 96.7%; Score 145; DB 5; Length 346;
Best Local Similarity 96.3%; Pred. No. 6.4e-13;
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASDYILGWFGGVPHEKKEEN 27
Db |||||:|||||:|||||:|||||:|||||
187 AKEASYNILGWFGGVPHEKKEEN 213

RESULT 11
Q9GS22 PRELIMINARY; PRT; 371 AA.
AC Q9GS22;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Merozoite surface protein 6.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21106479; PubMed=11166390;
RA Trucco C., Fernandez-Reyes D., Howell S., Stafford W.H.,
RA Scott-Finnigan T.J., Grainger M., Ogun S.A., Taylor W.R., Holder A.A.;
RT "The merozoite surface protein 6 gene codes for a 36 kDa protein
RT associated with the Plasmodium falciparum merozoite surface protein-1
RT complex.";
DR EMBL; AY007721; AAG15394.1; -.
DR InterPro; IPR001680; WD40.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
SQ SEQUENCE 371 AA; 42275 MW; 4788BC1FAF965CD6 CRC64;

Query Match 44.3%; Score 66.5; DB 5; Length 371;
Best Local Similarity 63.6%; Pred. No. 0.23;
Matches 14; Conservative 0; Mismatches 3; Indels 5; Gaps 1;

QY 10 ILGWFGGVP-----EHKKEE 26
Db |||||:|||||:|||||:|||||:|||||
182 ILGWFGGAPQNGAAEDKTE 203

RESULT 12
Q8IJ54 PRELIMINARY; PRT; 371 AA.
AC Q8IJ54;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Merozoite surface protein 6.
GN PF10_0346.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;

[1]
RN SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AE014834; AAN35543.1; -.
DR InterPro; IPR001680; WD40.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
SQ SEQUENCE 371 AA; 42275 MW; 4788BC1FAF965CD6 CRC64;

Query Match 44.3%; Score 66.5; DB 5; Length 371;
Best Local Similarity 63.6%; Pred. No. 0.23;
Matches 14; Conservative 0; Mismatches 3; Indels 5; Gaps 1;

QY 10 ILGWFGGVP-----EHKKEE 26
Db |||||:|||||:|||||:|||||:|||||
182 ILGWFGGAPQNGAAEDKTE 203

RESULT 13
Q8IJ48 PRELIMINARY; PRT; 405 AA.
AC Q8IJ48;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF10_0352.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AE014834; AAN35549.1; -.
DR Hypothetical protein.
SQ SEQUENCE 405 AA; 46729 MW; 5C6BC2106B94F5FD CRC64;

Query Match 44.0%; Score 66; DB 5; Length 405;
Best Local Similarity 47.6%; Pred. No. 0.3;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 7 YDYILGWFGGVPHEKKEEN 27
Db |:::|||||:|:|:|:|:|:|
232 YNHFAWEIGGAPTKYKPN 252

RESULT 14
O69947

```

```

AC Q7WR42;
01-OCT-2003 (TrEMBLrel. 25, Created)
01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12).
OS ASPS OR TLS OR B90115
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacterii; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
NCBI_TaxID=518;
[1]_SEQUENCE FROM N.A.
RC STRAIN=B50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Harris S.D., Mungall K.L.,
RA Cardeno-Tarraga A.M., Temple L., James K., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640437; CAE30616.1; -.
KW Aminoacyl-tRNA synthetase; Ligase; Complete proteome.
SQ SEQUENCE 596 AA; 67185 MW; F74DE904F8DDA448 CRC64;

Query Match      38.7%; Score 58; DB 16; Length 596;
Best Local Similarity 54.2%; Pred. NO. 7;
Matches 13; Conservative 4; Mismatches 5; Indels 2; Gaps 2;

QY      4 ASSYDYL-GWEFGGG-VPEHKKK 25
DB      | : || : ||| ||| | | : ||
        478 AKAYDWLVNGWEIGGSGVRIHREE 501

Search completed: September 22, 2004, 18:09:07
Job time : 30.562 secs

```

W. E. B. DUBOIS

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:47 ; Search time 29.6198 Seconds
(without alignments)
298.263 Million cell updates/sec

Title: US-10-774-602-13
Perfect score: 147
Sequence: 1 PEHKKENMLSHLYVSSKOKENISKEND 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

1 number of hits satisfying chosen parameters: 1017041

Maximum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147	100.0	329	5 Q9NFV9	Q9NFV9 plasmodium
2	147	100.0	346	5 Q9UG00	Q9UG00 plasmodium
3	147	100.0	354	5 Q25995	Q25995 plasmodium
4	147	100.0	354	5 Q81J55	Q81J55 plasmodium
5	147	100.0	361	5 Q95P15	Q95P15 plasmodium
6	147	100.0	379	5 Q25705	Q25705 plasmodium
7	147	100.0	379	5 Q25706	Q25706 plasmodium
8	147	100.0	379	5 Q9UEC4	Q9UEC4 plasmodium
9	147	100.0	380	5 Q26019	Q26019 plasmodium
10	143	97.3	64	5 Q15762	Q15762 plasmodium
11	70.5	48.0	405	5 Q81J48	Q81J48 plasmodium
12	64	43.5	602	5 Q81B95	Q81B95 plasmodium
13	62	42.2	424	5 Q81J53	Q81J53 plasmodium
14	58	39.5	7170	5 Q81L30	Q81L30 plasmodium
15	57	38.8	2249	5 Q81IY8	Q81IY8 plasmodium
16	55	37.4	425	6 Q95221	Q95221 oryctolagus

17	55	37.4	448	11	Q80V72	mus musculus
18	55	37.4	524	17	Q8TR11	Q8TR11 methanosarc
19	55	37.4	527	11	Q8RON3	Q8RON3 mus musculus
20	55	37.4	551	11	Q811K1	Q811K1 mus musculus
21	55	37.4	1056	11	Q8COP7	Q8COP7 mus musculus
22	54.5	37.1	116	12	Q80IM0	Q80IM0 eupatorium
23	54.5	37.1	584	5	Q8IDV8	Q8IDV8 plasmodium
24	54.5	37.1	1155	10	Q9FX38	Q9FX38 arabidopsis
25	54.5	37.1	3063	5	Q81284	Q81284 plasmodium
26	54	36.7	214	10	Q9AX37	Q9AX37 oryza sativ
27	54	36.7	371	5	Q9GS22	Q9GS22 plasmodium
28	54	36.7	371	5	Q81J54	Q81J54 plasmodium
29	54	36.7	839	9	Q64076	Q64076 bacterioph
30	54	36.7	839	16	Q31945	Q31945 bacillus su
31	54	36.7	983	5	Q8SSJ2	Q8SSJ2 encephalito
32	53	36.1	92	10	Q9LJ31	Q9LJ31 oryza sativ
33	53	36.1	351	4	Q9UFU6	Q9UFU6 homo sapien
34	53	36.1	426	4	Q9UHH9	Q9UHH9 homo sapien
35	53	36.1	426	4	Q9H4P7	Q9H4P7 homo sapien
36	53	36.1	2418	5	Q81IR9	Q81IR9 plasmodium
37	53	36.1	2625	5	Q8MMZ9	Q8MMZ9 dictyosteli
38	53	36.1	3242	5	Q9V4N8	Q9V4N8 drosophila
39	52.5	35.7	600	5	Q77355	Q77355 plasmodium
40	52.5	35.7	1079	10	Q93XF7	Q93XF7 zea mays (m
41	52.5	35.7	1132	5	Q815C3	Q815C3 plasmodium
42	52.5	35.7	1435	5	Q81LL4	Q81LL4 plasmodium
43	52.5	35.7	1883	5	Q81LI0	Q81LI0 plasmodium
44	52	35.4	63	16	Q81RM5	Q81RM5 bacillus an
45	52	35.4	315	5	Q97304	Q97304 plasmodium

ALIGNMENTS

RESULT 1

Q9NFV9 ID Q9NFV9 PRELIMINARY; PRT; 329 AA.
AC Q9NFV9; 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN MSP3.
OS Plasmodium falciparum (isolate 7G8).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7G8;
RX MEDLINE=20416497; PubMed=10960178;
RA Okenu D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in
Plasmodium reichenowi and Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 109:185-188(2000).
DR EMBL; AJ252287; CAB85901.1; -
FT NON_TER 1
FT NON_TER 329
SQ SEQUENCE 329 AA; 36916 MW; C5B045DB5E21A159 CRC64;

Query Match 100.0%; Score 147; DB 5; Length 329;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKENMLSHLYVSSKOKENISKEND 28
|
Db 193 PEHKKENMLSHLYVSSKOKENISKEND 220
|

RESULT 2

Q9UG00 ID Q9UG00 PRELIMINARY; PRT; 346 AA.
AC Q9UG00; 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Merozoite surface protein 3 (Fragment).
 GN MSP3.
 OS Plasmodium reichenowi.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5854;
 RN [1]_
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20416497; PubMed=10960178;
 RA Okenu D.M.N., Thomas A.W., Conway D.J.;
 RT "Allelic lineages of the merozoite surface protein 3 gene in
 RT Plasmodium reichenowi and Plasmodium falciparum";
 RL Mol. Biochem. Parasitol. 109:185-188 (2000).
 DR EMBL, AJ252286; CAB65754.1; -.
 FT NON_TER 1
 FT NON_TER 346
 SQ SEQUENCE 346 AA; 39127 MW; A804B96BDFAF010 CRC64;
 Query Match 100.0%; Score 147; DB 5; Length 346;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
 Db 206 PEHKKEENMLSHLYVSSKDKENISKEND 233
 RESULT 3
 Q25995
 ID Q25995 PRELIMINARY; PRT; 354 AA.
 AC Q25995;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Antigen.
 OS Plasmodium falciparum (isolate NF54).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5843;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=NF54;
 RC MEDLINE=95198774; PubMed=7891748;
 RA McColl D.J., Silva A., Foley M., Kun J.F., Favaloro J.M.,
 RA Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;
 RT "Molecular variation in a novel polymorphic antigen associated with
 RT Plasmodium falciparum merozoites";
 RL Mol. Biochem. Parasitol. 68:53-67 (1994).
 DR EMBL, L28825; AAC09377.1; -.
 SQ SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;
 Query Match 100.0%; Score 147; DB 5; Length 354;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
 Db 203 PEHKKEENMLSHLYVSSKDKENISKEND 230
 RESULT 4
 Q81J55
 ID Q81J55 PRELIMINARY; PRT; 354 AA.
 AC Q81J55;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)

DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Merozoite surface protein 3.
 GN PF10_0345.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]_
 RP SEQUENCE FROM N.A.
 RX STRAIN=3D7;
 RC MEDLINE=22255705; PubMed=12368864;
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
 RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
 RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
 RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrell B.;
 RT "Genome sequence of the human malaria parasite Plasmodium
 RT falciparum";
 RL Nature 419:498-511 (2002).
 DR EMBL, AE014834; AAN35542.1; -.
 SQ SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;
 Query Match 100.0%; Score 147; DB 5; Length 354;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
 Db 203 PEHKKEENMLSHLYVSSKDKENISKEND 230
 RESULT 5
 Q95P15
 ID Q95P15 PRELIMINARY; PRT; 361 AA.
 AC Q95P15;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Merozoite surface protein 3 (Fragment).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=FVO;
 RA Hisaeda H., Saul A., Long C.A., Miller L.H., Stowers A.W.;
 RT "Merozoite Surface Protein 3 and Protection Against Malaria in Aotus
 RT Monkeys";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AY044180; AAK94780.1; -.
 FT NON_TER 1
 FT NON_TER 361
 SQ SEQUENCE 361 AA; 41163 MW; 6127A3041587BA74 CRC64;
 Query Match 100.0%; Score 147; DB 5; Length 361;
 Best Local Similarity 100.0%; Pred. No. 2.2e-11;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
 Db 209 PEHKKEENMLSHLYVSSKDKENISKEND 236
 RESULT 6
 Q25705
 ID Q25705 PRELIMINARY; PRT; 379 AA.
 AC Q25705;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

```

RT gene. ";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RR EMBL; AF188190; AA04099.1; -.
RS
SQ SEQUENCE 379 AA; 43316 MW; C152A54E1F9D5F25 CRC64;

Query Match 100.0%; Score 147; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 2,3e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
Db 229 PEHKKEENMLSHLYVSSKDKENISKEND 256
|||||
|||||

RESULT 9
Q26019 PRELIMINARY; PRT; 380 AA.
ID AC Q26019;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Polymorphic antigen precursor.
DE Plasmodium falci-parum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC27;
RX MEDLINE=95198774; PubMed=7891748;
RA McColl D.J., Silva A., Foley M., Kun J.F., Favaloro J.M.,
RA Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;
RA "Molecular variation in a novel polymorphic antigen associated with
RA Plasmodium falci-parum merozoites.";
RT Mol. Biochem. Parasitol. 90:21-31(1997).
RL Mol. Biochem. Parasitol. 68:53-67(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FC27;
RX MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RA "Conservation of structural motifs and antigenic diversity in the
RT Plasmodium falci-parum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; L07944; AAC09378.1; -.
DR PDB; 1PSM; 07-FEB-95.
KW SIGNAL.
FT SIGNAL.
FT CHAIN 1 25 POTENTIAL.
FT CHAIN 26 380 POLYMORPHIC ANTIGEN.
SQ SEQUENCE 380 AA; 43290 MW; 0986CA1393094CA2 CRC64;

Query Match 100.0%; Score 147; DB 5; Length 380;
Best Local Similarity 100.0%; Pred. No. 2,3e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
Db 230 PEHKKEENMLSHLYVSSKDKENISKEND 257
|||||
|||||

RESULT 10
Q15762 PRELIMINARY; PRT; 64 AA.
ID AC Q15762;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
DE MSP-3.
OS Plasmodium falci-parum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=5843;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NF54;

```

```

MEDLINE=94348085; PubMed=8068948;
RX Oeuvray C., Bouharoun-Tayoun H., Gras-Masse H., Bottius E., Kaidoh T.,
RA Akawa M., Filgueira M.-C., Tartar A., Druilhe P.;
RT "Merozoite surface protein-3: a malaria protein inducing antibodies
RT that promote Plasmodium falciparum killing by cooperation with blood
RT monocytes.";
RL Blood 84:1594-1602(1994).
RN [2]
RC SEQUENCE FROM N.A.
RD STRAIN=NF54;
RA Oeuvray C., Bouharoun-Tayoun H., Gras-Masse H., Bottius E., Kaidoh T.,
RA Akawa M., Filgueira M.-C., Tartar A., Druilhe P.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF024624; AA81610.1; -.
FT NON TER 1
FT NON TER 64
FT NON TER 64
SQ SEQUENCE 64 AA; 7326 MW; 02E0ABF614C15A3B CRC64;

Query Match 97.3%; Score 143; DB 5; Length 64;
Best Local Similarity 96.4%; Pred. No. 1.4e-11;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PEHKKEENMLSHLYVSSKOKENISKEND 28
Db 37 PEHKKEENMLSHLYVSSKOKENISKENE 64

RESULT 11
Q81J48 PRELIMINARY; PRT; 405 AA.
ID Q81J48
AC Q81J48;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF10_0352.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Perteira M., Allen J., Salengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McPadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
RD EMBL; AE014834; AAN35549.1; -.
RW Hypothetical protein.
KW
SQ SEQUENCE 405 AA; 46729 MW; 5C6BC2106E94F5FD CRC64;

Query Match 48.0%; Score 70.5; DB 5; Length 405;
Best Local Similarity 48.3%; Pred. No. 0.31;
Matches 14; Conservative 8; Mismatches 6; Indels 1; Gaps 1;

Qy 1 PEHKKEEN-MLSHLYVSSKOKENISKEND 28
Db 249 PENNKNDNILLEHVKITSWDKEDIKENE 277

RESULT 12
Q81B95 PRELIMINARY; PRT; 602 AA.
ID Q81B95
AC Q81B95;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)

```

```

RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrill B.G;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
RR EMBL; AL034557; CAD49133.1; -.
DR EMBL; AL034557; CAD49133.1; -.
KW Hypothetical protein.
SQ SEQUENCE 2249 AA; 271088 MW; 6EF04691C22F8797 CRC64;

Query Match 38.8%; Score 57; DB 5; Length 2249;
Best Local Similarity 47.8%; Pred. No. 99;
Matches 11; Conservative 5; Mismatches 7; Indels 0;

QY 3 HKKEENMLSHLYVSSKDKENISK 25
      ||| ||| |||
      :|: |||
      :|: |||

Db 882 HKTESNNLTHINQKEDKKNISK 904

Search completed: September 22, 2004, 18:09:08
Job time : 30.6198 secs

```

100